FEB 4 1942 LIBRARY

The Journal of Experimental Education

A periodical report of scientific investigations relating to child development, curriculum, learning, teaching, supervision, measurements, statistics, and experimental techniques.

Volume X | December, 1941 | Number 2

CHILD DEVELOPMENT

CONTENTS

PAGE

Seasonal Variations in I.Q.: *Harold E. Jones* 91

Children's Interest in Reading the Comics: *Paul Witty* 100

Reading the Comics—A Comparative Study: *Paul Witty* 105

Words and the Comics: *Robert L. Thorndike* 110

An Approach to the Measurement of Insecurity: *Elinor Prichard and Ralph H. Ojemann* 114

Studies of Elementary School Classes in Action—II. Pupil Participation and Aspects of Pupil–Teacher Relationships: *Arthur T. Jersild, Bernard Goldman, Catherine L. Jersild, John J. Loftus* 119

Observations of the Behavior of Children in Activity and Control Schools: *Robert L. Thorndike, John J. Loftus, Bernard Goldman* 138

Observations of Excursions in Activity and Control Schools: *Robert L. Thorndike, John J. Loftus, Bernard Goldman* 146

Language Development of Children in Two Types of Rural Environments: *G. M. Worbois* 150

$5.00 a Year | Published Quarterly | $1.50 a Copy

Edited and Published by A. S. Barr, Professor of Education, University of Wisconsin, Madison, Wisconsin.

Entered as second-class matter October 17, 1938 at the post office at Madison, Wisconsin, under the act of March 3, 1879.

EDITORIAL BOARD

A. S. Barr, Chairman, Professor of Education, University of Wisconsin, Madison, Wis.

Carter V. Good, Professor of Education, University of Cincinnati, Cincinnati, Ohio. Editorially responsible for materials on learning, teaching and supervision, published each September.

J. Wayne Wrightstone, Assistant Director, Bureau of Reference, Research and Statistics, Board of Education of the City of New York, 110 Livingston Street, Brooklyn, New York. Editorially responsible for materials on curriculum construction, published each June.

Edward E. Cureton, Senior Educational Statistician, U. S. Office of Education, Washington, D. C. Editorially responsible for materials on measurements, statistics, and methods of experimental research, published each March.

Arthur T. Jersild, Professor of Education, Teachers College, Columbia University, New York City. Editorially responsible for materials on child welfare, guidance, and development, published each December.

CONTRIBUTING EDITORS

Gilbert L. Betts, Supervisor of Graduate Research in Education, Colorado State College, Fort Collins, Colorado.

William A. Brownell, Professor of Educational Psychology, Duke University, Durham, North Carolina.

Leo J. Brueckner, Professor of Education, University of Minnesota, Minneapolis, Minnesota.

Oscar K. Buros, Associate Professor of Education, Rutgers University, New Brunswick, New Jersey.

Otis W. Caldwell, General Secretary, The American Association for the Advancement of Science, Boyce Thompson Institute for Plant Research, Yonkers, New York.

Leslie L. Chisholm, Associate Professor of Education, State College of Washington, Pullman, Washington.

Herbert S. Conrad, Associate Professor of Psychology, College of Agriculture, and Research Associate, Institute of Child Welfare, University of California, Berkeley, California.

Stephen M. Corey, Professor of Educational Psychology, University of Chicago, Chicago, Illinois.

Robert A. Davis, Professor of Education, Director of Bureau of Educational Research, University of Colorado, Boulder, Colorado.

Harl R. Douglass, Director of College of Education, University of Colorado, Boulder, Colorado.

Jack W. Dunlap, Associate Professor of Educational Psychology, University of Rochester, Rochester, New York.

Harold A. Edgerton, Assistant Professor of Psychology, Ohio State University, Columbus, Ohio.

Alvin C. Eurich, Professor of Education, Stanford University, Stanford University, California.

John C. Flanagan, Assistant Chief, Research Section, Medical Division, Office, Chief of the Air Corps, War Department, Washington, D. C.

Kai Jensen, Associate Professor of Education, University of Wisconsin, Madison, Wisconsin.

Harold E. Jones, Professor of Psychology, Director of Research, Institute of Child Welfare, University of California, Berkeley, California.

Noel Keys, Professor of Education, University of California, Berkeley, California.

Edward A. Lincoln, Consulting Psychologist, Halifax, Massachusetts.

T. E. Newland, Pennsylvania State Department of Education, Chief of Special Education, Harrisburg, Pennsylvania.

C. W. Odell, Associate Professor of Education, University of Illinois, Urbana, Illinois.

Willard C. Olson, Professor of Education, Director of Research in Child Development, University of Michigan, Ann Arbor, Michigan.

W. E. Peik, Dean and Professor of Education, University of Minnesota, Minneapolis, Minnesota.

S. L. Pressey, Professor of Educational Psychology, Ohio State University, Columbus, Ohio.

Clarence E. Ragsdale, Associate Professor of Education, University of Wisconsin, Madison, Wisconsin.

H. H. Remmers, Director, Division of Educational Reference, Professor of Education and Psychology, Purdue University, Lafayette, Indiana.

Henry D. Rinsland, Professor of Education and Director of Educational Research, The University of Oklahoma, Norman, Oklahoma.

Robert T. Rock, Jr., Professor of Psychology, Head of Dept. of Psychology, Graduate School, Fordham University, New York City.

G. M. Ruch, Chief of Research and Statistical Service, U. S. Office of Education, Washington, D. C.

P. J. Rulon, Assistant Professor of Education, Graduate School of Education, Harvard University, Cambridge, Massachusetts.

Douglas E. Scates, Associate Professor of Education, Duke University, Durham, North Carolina.

David Segel, Educational Consultant, Specialist in Tests and Measurements, Federal Security Agency, U. S. Office of Education, Washington, D. C.

Paul W. Terry, Professor of Educational Psychology, University of Alabama, University, Alabama.

Helen Thompson, Clinic of Child Development, Research Associate, Yale University, New Haven, Connecticut.

Robert L. Thorndike, Associate Professor of Education, Teachers College, Columbia University, New York City.

Herbert A. Toops, Professor of Psychology, Ohio State University, Columbus, Ohio.

T. L. Torgerson, Professor of Education, University of Wisconsin, Madison, Wisconsin.

Helen M. Walker, Professor of Education, Teachers College. Columbia University, New York City.

Beth L. Wellman, Professor of Psychology, Child Welfare Research Station, State University of Iowa, Iowa City, Iowa.

Guy M. Wilson, Professor of Education, Boston University, Boston, Massachusetts.

Paul A. Witty, Professor of Education, Director of Psycho-Educational Clinic, School of Education, Northwestern University, Evanston, Illinois.

Ernest R. Wood, Professor of Education, New York University, New York City.

DEMOCRAT PRINTING COMPANY
MADISON, WISCONSIN

Journal of Experimental Education

Volume X | December, 1941 | Number 2

SEASONAL VARIATIONS IN IQ

Harold E. Jones[1]
University of California
Institute of Child Welfare

In one of the earliest of the University of Iowa reports on the effect of preschool attendance upon mental growth, Wellman (11) called attention to a favorable IQ change among nursery school children when tested in the fall and retested approximately six months later; a similar change was not observed when the spring tests were compared with subsequent measurements after a summer interval. These results were based chiefly on the Kuhlmann revision of the Binet for ages up to three-and-a-half years, and on the Stanford revision at older ages. Among 45 children first tested in the fall, an average gain of 9.9 IQ points was recorded in a spring retest; a small change of less than a point occurred in the ensuing summer, but the second winter interval was marked by a further gain of five points. When the data were combined for all cases with at least one fall and one spring testing, an average IQ gain of 7.8 was shown in fall-to-spring records for 310 pairs of tests; on the other hand, from spring-to-fall a slight average loss occurred, in a total of 193 pairs of tests. The same material appeared in a further report (12) in which IQ's were transformed into percentile scores; the fall-to-spring gain was reported as statistically significant.

A number of criticisms have been leveled at the statistical procedures in these studies, particularly as to the methods used in combining results from different tests, the averaging of percentile scores, and the computation of reliability indices for averages in which the same individuals are represented repeatedly (10). These technical objections, however, do not succeed in overcoming such data as those presented in Table 1, in which Wellman's results are subdivided according to the order of testing.

Table I

Seasonal Changes in IQ

(From Wellman[1])

Tests	Fall-to-Spring N	Fall-to-Spring Av. gain	Spring-to-Fall N	Spring-to-Fall Av. gain
1st vs 2nd...	152	9	31	3
2nd vs 3rd...	39	7	100	—2
3rd vs 4th...	76	6	27	—5
4th vs 5th...	18	8	28	—1
5th vs 6th...	20	4	6	—4

[1] In the absence of tabular data in the original study, the values in this table were determined, to the nearest integer, from Figure 6 in reference 11.

The reader will wish to know whether the fall-to-spring gain, taking account of sign, is reliably greater than the spring-to-fall gain, for the several comparisons in Table 1. This seems doubtful, and no reliability data are furnished. It would be surprising, however, if consistent differences of this order were to appear solely as a result of sampling errors or of errors of measurement. The differences, while perhaps not fully convincing, nevertheless cannot be lightly dismissed.

Wellman's explanation of her findings has emphasized the effects of nursery school attendance. If we allow for vacation periods and absences, it is probable that the fall-to-spring group had an average of about four months of school attendance, while the spring-to-fall group had an average of between six weeks and three months, according to whether or not they attended a summer session. Could so small a difference in attendance be responsible for the IQ differences which have been noted? In a subsequent article, Wellman (13) supported her argument by comparing seasonal changes for (a) 34 children not in preschool and (b) matched records from 404 preschool cases;

[1] The computations for this study were made by Ada P. Jorgensen and Luana B. Schrader, with the assistance of the personnel of Works Progress Administration Official Project No. 665-08-3-30. Acknowledgment is due to Dr. H. S. Conrad for suggestions concerning statistical procedures.

the matching was on the basis of age, allowing a variation of one year, and of IQ, allowing a variation of 10 points. From fall to spring, the preschool data showed a gain from an average IQ of 120.1 to an average IQ of 127.1. The non-preschool cases showed a loss in the same period (120 to 116.1).

Recognizing the need for a larger number of non-preschool cases to use in this comparison, Lodge (7) conducted a study in Cleveland, Ohio, among 171 children who were tested with the Stanford–Binet four times at six-month intervals, from three-and-a-half to five years of age.

Table 2 presents a portion of Lodge's results, for months most comparable to Wellman's study.

TABLE II

SEASONAL CHANGES IN IQ

(Cleveland, Ohio. Adapted from Lodge (7))

Ages	Fall-to-Spring		Spring-to-Fall	
	N	IQ gain	N	IQ gain
3½ to 4	27–28	7.2	28–29	—1.7
4 to 4½	64–61	2.2	61–64	— .4
4½ to 5	29–29	2.4	64–64	—1.2

The "fall" tests, in the above table were given at birthdays or half-birthdays in August, September or October; the "spring" tests in February, March and April.

Although no differences are statistically significant, in each instance they are consistent with the earlier data of Wellman. Impressed by this agreement, he concluded,

> ". . . The data demonstrate two points but do not differentiate between them: either early spring is the most favorable time for testing and early autumn the least favorable; or mental growth as reflected in the test scores is accelerated in late autumn and retarded in late spring. These results are not inconsistent with previous studies that have tended to support the view as to constancy of the IQ. The present findings instead of contradicting the notion of constancy actually indicate that the variation ordinarily expected may be somewhat reduced if the seasonal factor is held constant—if the children are tested at the same time each year."

Lodge's sample, of superior socio-economic status, was described as not receiving the benefits of preschool attendance as a group, and as not being selected on the basis of preschool experience. The reader is left with the uneasy feeling that an unstated proportion of these cases consisted of nursery school children; while Lodge is inclined to reject an explanation of the results in terms of nursery school influence, his data would be more readily interpretable if the preschool and non-preschool cases were clearly segregated.

The first part of the present study deals with 469 pairs of tests[1] of children between two and five years of age, in attendance at the Institute of Child Welfare nursery school (Berkeley, California). Two hundred seventy-four of the test pairs involve a testing in the fall compared with a retest approximately six months later in the spring; the mean IQ gain was 4.6 points. The remaining 195 test pairs involve a spring-to-fall sequence, and show a mean IQ gain of 1.6 points. While these results are suggestive, it is not feasible to compute critical ratios because the same individuals are represented in more than one pair of tests. To control this factor, and also to allow for differential effects which may occur in relation to the amount of previous test practice, the data have been subdivided as shown in Table 3 and in Figure 1.

As is usually found, Table 3 exhibits IQ gains which are greatest in the first as compared with the second test and which tend to diminish thereafter; this is to be expected if the initial experience with tests leads to an improvement through adaptation.

[1] See reference 6, p. 214, and footnote on p. 209, for a discussion of the tests used and the method for transforming raw scores for each test into comparable measures. Although the various tests involved considerable overlapping as to content, the precaution was taken to distribute tests in such a way that seasonal differences could not be attributed to variations in the instruments of measurement.

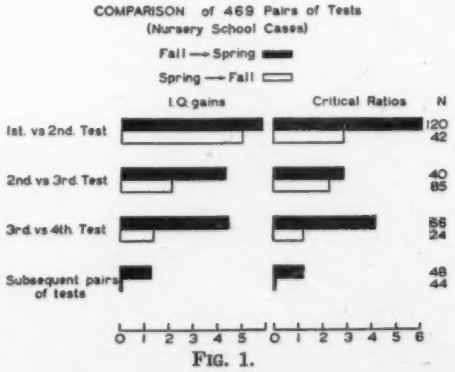

FIG. 1.

TABLE III

SEASONAL CHANGES IN IQ

(University of California, Nursery School Cases)

	Fall-to-Spring			Spring-to-Fall		
	N	IQ Gain	C. R.[1]	N	IQ Gain	C. R.
1st vs. 2nd Test	120	5.9	6.4	42	5.1	2.9
2nd vs. 3rd Test	40	4.4	2.9	85	2.2	2.3
3rd vs. 4th Test	66	4.5	4.2	24	1.4	1.2
Pairs of Subsequent Tests	48	1.3	1.2	44	.1	.1

Mean C. A. at the 1st test: Fall-to-Spring, 35.6 Months
Spring-to-Fall, 34.3 Months

[1] In testing the significance of the difference the formula used is

$$\sigma_{diff} = \sqrt{\sigma_{m_1}^2 + \sigma_{m_2}^2 - 2r\sigma_{m_1}\sigma_{m_2}}$$

The fall-to-spring gains are significant for the first through the fourth test, while the spring-to-fall gains are significant only for the first to the second. The differences in the critical ratios are of course small, but as in the Wellman and Lodge data a consistent tendency can be noted for greater gains to occur during the fall-to-spring interval. This is not to be explained in terms of sampling differences, since substantially the same groups are involved in rotation in the first two fall-spring and spring-fall pairs of comparisons.

If the question is raised as to the role of nursery school attendance in these findings, two preliminary approaches can be made to an answer. It is possible to subdivide all spring-to-fall groups in terms of whether or not they attended nursery school during six weeks of the summer interval. The results are given in Table 4.

With these smaller groups, the IQ changes, although generally positive, are not statistically significant, nor do they generally favor the nursery school group. It should be pointed out that the "comparison group" has had nursery school experience in the regular session; this would tend to reduce differences that might otherwise be found. Moreover, since the policy has been to test children near their annual and six-month birthdays, rather than precisely at the end of the spring term and the beginning of the fall term, some nursery school attendance is included in all spring-to-fall intervals. More suitable for the present purpose is the comparison of seriatim records for children who have left nursery school, with those who have continued in school. Table 5 gives the mean IQs for 58 children who received four successive tests, in comparison with 13 who received the same number of tests but left school after the third test.

Table 6 presents similar data for 30 nursery school children receiving five con-

TABLE IV

IQ CHANGES (SPRING-TO-FALL) IN RELATION TO SUMMER SESSION NURSERY SCHOOL ATTENDANCE

	Nursery School Group			Comparison Group		
	N	IQ Gain	C. R.	N	IQ Gain	C. R.
First vs. Second Test	18	6.4	2.6	24	4.6	2.0
Second vs. Third Test	39	2.2	1.3	46	1.8	1.6
Pairs of Subsequent Tests	22	—.6	—.6	46	1.1	1.1

TABLE V

MEAN IQS FOR TWO GROUPS WITH FOUR CONSECUTIVE TESTS

Test and Semester:	N	1	2	3	4
Nursery School Group	58	109.0	117.0	120.2	123.8
Group Leaving School after Third Semester	13	109.6	114.3	115.9	121.1

TABLE VI

MEAN IQS FOR TWO GROUPS WITH FIVE CONSECUTIVE TESTS

Test and Semester:	N	1	2	3	4	5
Nursery School Group	30	108.6	117.1	121.7	122.5	123.1
Group Leaving School after Third Semester	9	111.0	115.1	118.0	123.7	123.8

secutive tests, and for 9 children leaving school after the third test.

Tables 5 and 6 do not present data with reference to seasonal variation. They tend to show, however, that IQ gains occur in the same way whether or not children are in nursery school. Other evidence analyzed in reference (6) has indicated that IQ gains, when they occur in this sample, must be attributed to factors other than nursery school attendance.

We may now turn to evidence on seasonal variation in groups which have had no nursery school attendance at any time. The Guidance Study of the Institute of Child Welfare consists of 126 children who were tested, under the direction of Dr. J. W. Macfarlane, with California Preschool Schedules I or II[1] at 21 and 24 months (average age 23 months) and at 30, 36, 42, and 48 months. From this sample, which was originally selected as representative of children born in Berkeley during 1928 and the first half of 1929, cases were taken who had their first tests in September, October, and November, and a second test at the age of 30 months in the following spring. Similarly, cases were selected who had their first tests in March, April or May, and a second test in the fall. No children with nursery school experience were included. Table 7 presents IQ changes for this sample, which on the whole are smaller than for the more highly selected sample in Table 3. None of the changes are statistically significant, except for the fall-to-spring group between the ages of 42 and 48 months.[2] Caution must be observed in drawing conclusions from this material as to seasonal variations; in three out of four comparisons the obtained change is greater for the fall-to-spring group, but in only one of these is the difference reliably greater. (The C. R. is 3.19 for the difference in gains at 42 to 48 months).

[1] See reference 8 for a description of these tests.

[2] No critical ratio was computed for the "all ages" comparison, since in this the same individuals occur more than once.

TABLE VII

SEASONAL CHANGES IN IQ

(Guidance Study, Non-Nursery School Cases)

Ages	Fall to Spring N	Fall to Spring IQ Gains	Spring to Fall N	Spring to Fall IQ Gains
23 to 30 Months	16	2.5	31	1.5
30 to 36 Months	33	2.3	18	.8
36 to 42 Months	17	—1.8	25	.2
42 to 48 Months	22	8.0	13	.8
All Ages	88	3.0	87	.9

One other source of evidence is available, from a sample of Berkeley children tested from early infancy by Dr. Nancy Bayley (1). Twenty-four of these cases were tested in the spring months (March, April or May) at the age of either 24 or 30 months, and six months later (at the ages of 30 and 36 months) were tested in September, October or November. The tests were repeated at six-month intervals to the ages of 54 and 60 months. The California Preschool Schedules I or II were used, and the raw scores transformed into sigma scores (see reference (1)) to eliminate effects of test practice. Since all of these cases had been tested at the Institute from early infancy under the same external conditions and by the same examiner, it may be assumed that the influence of extraneous variables in the procedure was reduced to a minimum. Table 8 presents results from this sample, based on a total of 101 pairs of tests.

Four comparisons are possible, in Table 8, for sequential spring-fall and fall-spring changes; in each of these the difference is in favor of the fall-spring period, the value of the difference being on an average .23 of the S. D. of the distribution of the total group from which these cases were drawn. The average sigma score for two spring testings is .115, for three fall testings — .12.

TABLE VIII

MENTAL SIGMA SCORE CHANGES IN RELATION TO SEASON

(Berkeley Growth Study, Non-Nursery School Cases)

	N	Change
Spring to Fall (24 to 30 mo. 30 to 36 mo.)	21	—.10
Fall to Spring (30 to 36 mo. 36 to 42 mo.)	21	+.01
Spring to Fall (36 to 42 mo. 42 to 48 mo.)	19	—.10
Fall to Spring (42 to 48 mo. 48 to 54 mo.)	19	+.22
Spring to Fall (48 to 54 mo. 54 to 60 mo.)	21	—.16

The foregoing differences are not statistically significant. However, in view of similar findings in other Berkeley samples, and the agreement of the California results with those previously reported from Iowa and Ohio, it is clear that we cannot dismiss the hypothesis of seasonal variation as wholly unfounded.

One limitation of the material thus far discussed is the size of the interval between tests; if, for example, climatic factors are assumed to have some effect upon mental functioning as expressed in intelligence tests, it would be helpful if a further analysis could be based on monthly rather than "fall-to-spring" vs. "spring-to-fall" retests. The literature on physical growth includes a number of references in which such an analysis has been made. With regard to weight, for example, Marshall (9) has summarized ten studies, eight of which reported a seasonal variation, with maximal growth in the fall and minimal changes in winter and spring. Representative of these is the investigation by Whitacre (14) who showed, for various white, negro and Mexican groups in Texas, an average weight increase in October amounting to from 11.3 per cent to 16.3 per cent of the yearly growth, while in April the increase was only from 1.4 per cent to 6.4 per cent of the yearly growth. Emerson (2) has concluded that such variations are not a basic expression of growth rhythms, but represent rather the influence of unfavorable conditions, such as infections (particularly respiratory infections) which have a greater incidence in the winter. It is also possible that conditions associated with a greater degree of outdoor life in the summer, with exposure to sunlight and perhaps nutritional differences, may have a stimulating effect upon growth. The physiological expression of such effects, as in body weight, would be noted only after a time lag of one or more months. It is conceivable that such factors would also exert an influence upon the central nervous system, either directly or as an outcome of improved general physical condition; the resulting changes in mental functioning might appear after a time lag not necessarily of the same duration as that observed for growth in body weight. Among children between the ages of two and seven years, Honzik and Jones (3) have shown a small degree of positive correlation between mental scores and physical measures (weight and height) and also between gains in mental scores and gains in physical measures. This relationship depends partly upon socio-economic factors, but is not wholly eliminated when a socio-economic index is partialled out. If the hypothesis is offered that the mental-physical relationship involves, in part, a seasonal reference, the data should be analyzed to determine whether the maximal seasonal change in IQ falls in some observable relation to the maximal changes which have been noted for physical measures.

In the absence of mental scores for tests administered at monthly intervals, a possible recourse is to compute the amount of IQ change for each month by means of a fractionated moving average.

Following is an example of this procedure. Thirty Guidance Study cases were tested in the month of September, and re-tested in March with an average gain of 3.0 IQ points. This gain was credited to each month in the period from September 15 to March 15, on the assumption of an even rate of growth within the six-month period. Similar calculations were made for cases tested in October and re-tested in April, etc.; since each month was represented repeatedly in these calculations, it was possible to determine an average gain for each month. The same procedure was used with Lodge's data, except that a somewhat coarser smoothing was made necessary by the fact that his results were reported not according to the month of

testing but for groups of three months. Because of the lack of detailed data, comparisons could not be made with the Wellman study.

Table 9, and, for the Berkeley data, Figure 2, exhibit the result of this analysis. The similarity between the two arrays of monthly values can be expressed in terms of a rank-order correlation of .87.

FIG. 2.

TABLE IX

IQ GAINS BY MONTH

	Berkeley Data	Cleveland Data
January	.36	.14
February	.31	.01
March	.23	—.12
April	.21	—.24
May	.16	—.26
June	.13	—.16
July	.17	—.06
August	.23	.07
September	.30	.22
October	.32	.39
November	.38	.42
December	.41	.27

The Cleveland maximum is seen to be in October and November, the Berkeley maximum in November and December. Likewise, the Cleveland minimum, dipping into negative values, is in April and May, the Berkeley minimum in May and June. If the slight displacements between the two series have any significance as regards climatic effects, one might speculate that the greater inter-monthly range in Cleveland is related to a greater range of climatic differences; that the more rapid decrement in Cleveland following the month of November is related to more adverse conditions in the late fall and early winter; and that the Berkeley lag in reaching a peak value is related to the fact that the peak of summer weather and the onset of contrasting "fall" or "winter" weather are definitely later than in eastern localities.

Lest this seem a heavy weight of conjecture resting upon what are, after all, small differences, it can be pointed out that the use of moving averages has necessarily resulted in lowering the values for maximum change and raising the values for minimum change. In the Berkeley data the results indicate (on the average) about three times the amount of change in November or December as in June. This is almost certainly an understatement of the actual difference between these months.

In computing Table 9, since no monthly differentials in IQ change could be assumed in advance, the original smoothing was necessarily conducted on the assumption of a constant rate of change within any given six-month period. But the values obtained in Table 9 in themselves make it possible to correct this assumption by first approximations to the true differentials; when this is done, a monthly weighting can be applied which will result in adjusted values constituting a second approximation. For the Berkeley data this second approximation gives an average IQ gain of .44, .51 and .49 for November, December and January respectively; and of .10, .07, and .09 for May, June and July respectively.

The statement should be added that no statistical manipulation can provide a fully adequate substitute for short-interval re-tests; one purpose in reporting this analysis is to stimulate investigators, in different geographic areas, to conduct such re-tests under conditions which will adequately control age and practice factors.

SUMMARY

1. From Institute of Child Welfare Nursery School records, 469 pairs of intelligence tests were examined with reference to seasonal variations. Controlling age and practice factors, it was found that IQ gains were

consistently although not reliably greater from fall-to-spring than from spring-to-fall. The average gain over the winter interval was 4.6 IQ points, over the summer interval 1.6.

2. Two types of analysis led to discounting the effects of nursery school training in producing gains: (a) for the spring-to-fall interval, no consistent IQ changes were noted in relation to summer session nursery school attendance; (b) cumulative testing of nursery school cases showed a similar course of IQ change whether or not they remained in nursery school.

3. 175 pairs of tests were examined from cumulative records of non-nursery school children with birthdays in fall or spring. At ages two to four years IQ gains were greater in fall-to-spring in three out of four comparisons, but were reliably different in only one comparison. The average IQ gain was 3.0 points fall-to-spring, .9 spring-to-fall.

4. 101 pairs of tests were examined from a third sample of non-nursery school children cumulatively tested between the ages of two and five years. Differences were not reliable but were consistently in favor of the fall-to-spring interval, the difference averaging about one-quarter of a standard deviation, in four comparisons.

5. From a series of 329 pairs of tests given at six-month intervals and representing all months of the year, the average IQ gain per month was computed. Months of greatest gain were November, December and January; months of least gain were May, June and July. A similar pattern of IQ change was obtained in an analysis of data from Cleveland, Ohio, reported by Lodge (7).

Because of the small magnitude and, for the most part, the statistical unreliability of the seasonal variations noted, it might appear justifiable to conclude that the evidence is negative as to the existence of such variations in mental scores. On the other hand, if the results of the present study are added to those of Wellman and Lodge, we now have available nineteen comparisons of fall-to-spring vs. spring-to-fall IQ changes, in children of preschool age (see Tables 1, 2, 3, 7, 8). These comparisons are based on a variety of preschool intelligence tests, and on five different samples in three widely separated localities. Eighteen of the nineteen comparisons show a greater gain over the winter interval than over the summer interval.

Discussion

Several alternative hypotheses can be considered in interpreting these results.

1. Hypothesis: The differences are due to a seasonal factor involving the manner in which the tests are administered and scored, or, in other words, a seasonal variation in the testers.

It is very difficult to eliminate this possibility. It is also difficult to take it very seriously, especially where results are based on a considerable number of expert examiners. Are we to assume, for example, that examiners are more lenient in January than in April (see Table 9), or that for some reason they are able to elicit better cooperation in November than in May?

2. Hypothesis: The effects noted, if real, are due to seasonal differences, of a transitory nature, in test *performance*. Such differences may rest upon some handicapping effect in spring and summer, some stimulating effect in the fall and early winter.

The influence of weather upon mental and physical efficiency has been discussed by Huntington (4, 5) with particular reference to such factors as mean temperature, changes in temperature, and the incidence of storms. In plotting curves for the output of factory operatives in Connecticut and Pittsburgh, the maximum was most frequently found in November; similar curves based on records from the southern states established a most frequent high point in December. These results would seem to be in accord with those reported by Lodge and in the present study respectively, but an explanation of the present data in meteorological terms (i. e., in terms of weather conditions during the month in which the test occurs) cannot be regarded as adequate.

Referring to Table 9, the effect of unfavorably low temperatures may possibly be seen in the Cleveland data, in the rapid decline from a maximum IQ gain in November; some additional factor, however, would be necessary to explain the continuation of the decline into April and May. In Berkeley, where summer and winter mean temperatures differ by only

about 12° F., one would expect little influence from this factor. The stimulating effect of storms, as noted by Huntington, might seem to provide a partial explanation for the Berkeley data, since storms are prevalent during the period of maximum IQ gains and absent during the period of minimum gains. The values in Table 9, however, show a very imperfect relationship to the incidence of storms, since the IQ gains tend to accelerate in the early autumn before the onset of storms, and fall off sharply in the late winter and early spring when storms are still frequent. Moreover, variations in this factor cannot explain the Iowa and Ohio results.

No other climatic influences or combinations of factors have been suggested which appear to parallel in any way the observed changes in mental scores, if these are conceived solely as representing temporary variations in efficiency. We are then led to the further alternative that the seasonal differences are due less to environmental conditions at the time of the test than to factors which have operated over a period of time during the test interval.

3. Hypothesis: The seasonal variations are related to intermediary differences in the child's own activity associated with season. These activities may in the winter more than in the summer involve practice on tasks similar to the content of intelligence tests.

While this hypothesis carries some weight, it is difficult to see why, in these terms, the maximal transfer occurs in November and December rather than at the end of winter, when the effects of such activities have had a longer time in which to accumulate.

4. Hypothesis: The results of this study indicate a seasonal variation in basic rates of mental growth, related to the seasonal changes previously established for growth in weight. The summer factors which favor and the winter factors which handicap mental and physical growth (see page 19–20) produce effects not immediately noted; these effects become manifest after a time lag which varies for different functions.

This is, of course, an environmental hypothesis. It remains to be determined whether the environmental effects associated with season are in any degree controllable; whether, for example, the favorable effects can be promoted, the unfavorable effects prevented or compensated. In any event, it is not assumed that the variations due to such factors play more than a very minor role in determining individual differences in the pattern of mental growth.

References

1. Bayley, N. "Mental Growth in Young Children," *Part II, 39th Yearbook, National Society for the Study of Education,* 1940, (Chap. II, pp. 11–47).
2. Emerson, H. "Seasonal Variation in Growth of School Children," *Journal American Medical Association,* LXXXIX, (1927), pp. 1326–1330.
3. Honzik, M. P. and Jones, H. E. "Mental-physical, Relationships During the Preschool Period," *Journal of Experimental Education,* VI, (1937), pp. 139–146.
4. Huntington, E. *Season of Birth: Its Relation to Human Abilities.* New York: Wiley, 1938.
5. Huntington, E. *Civilization and Climate.* New Haven: Yale University Press, 1924, 3rd. ed., pp. 455.
6. Jones, H. E. and Jorgensen, A. P. "Mental Growth as Related to Nursery School Attendance," *Part II, 39th Yearbook National Society for the Study of Education,* 1940, (Chap. XII, pp. 207–222).
7. Lodge, T. "Variation in Stanford–Binet IQs of Preschool Children According to the Months in which Examinations were Given," *Journal of Psychology,* VI, (1938), pp. 385–395.
8. Macfarlane, J. W. *Studies in Child Guidance. I. Methodology of Data Collection and Organization.* Monograph, Society for Research on Child Development, III, (1938), p. 254.
9. Marshall, E. L. "A Review of American Research on Seasonal Variation in Stature and Body Weight," *Journal of Pediatrics,* X, (1937), pp. 819–831.
10. McNemar, Q. *A Critical Examination of the University of Iowa Studies of Environmental Influence upon the IQ.* Psychological Bulletin, XXXVII, (1940), pp. 63–92.

11. Wellman, B. L. "Some New Bases for Interpretation of the IQ," *Journal of Genetic Psychology*, XLI, (1932), pp. 116–120.
12. Wellman, B. L. "The Effect of Preschool Attendance upon the IQ," *Journal of Experimental Education*, I, (1932), pp. 48–69.
13. Wellman, B. L. "Growth in Intelligence under Differing School Environments," *Journal of Experimental Education*, III, (1934), pp. 59–83.
14. Whitacre, J. *Growth in Height and Weight of Texas School Children*. Texas Agricultural Experimental Station, XLVII, (1934), pp. 122–124.

CHILDREN'S INTEREST IN READING THE COMICS

PAUL WITTY
Northwestern University

John, aged nine, approached me as I entered a school last spring: "Mr. Witty, would you like to rent a book?" Twenty comic magazines were displayed, and an attractive proposition was then set forth. For a very small sum, I could receive "Secret Service" or "Green Revenge"—o r i g i n a l, colored comics which John and his friends would make. A slightly larger sum would entitle me to a special limited edition, designed and tailored apparently to suit my personality.

I have observed over a considerable period of time various other manifestations of interest in the comics. In 1925–1927, Professor Lehman and I reported that "looking at the Sunday funny paper" was the play activity in which boys and girls of ages 8 to 15 engaged most frequently.* Interest in the comics has grown, it appears, during the ensuing period. One investigator states that 80,000,000 Americans now read the Sunday supplement containing comics. In another study it was found that the average youngster in grades IV, V, and VI stated that he read twenty-three comic strips "all the time".** Moreover, our newstands are surfeited with comic books and magazines.

Teachers and parents are becoming increasingly aware of the extent to which children read the comics. Reactions to this activity are varied. Some persons look upon the situation with apprehension or marked disfavor. For example, Sterling North's famous editorial in which he vigorously condemned the comics was favorably received and endorsed by many adults. In fact, some parent-teacher organizations promptly tabooed the sale of comics. In one locality in which this occurred, ingenious youngsters quickly developed "rental libraries"—their answer to their parents' hasty prohibition.

On the other hand, there are many adults who are undisturbed by the recent preoccupation of children with the comics. They look upon this activity as a product of many conditions and factors in our culture and assert that this reading habit simply parallels other contemporary developments such as an intense interest in highly exciting radio programs and movies. Through these media, adventure, excitement, and "humor" are offered the child with a minimum of effort and concern on his part.

In recent investigations, the comic strip, the movie, and the radio have been cited again and again as attaining high rank among the recipients of children's interest. Children sometimes assign first place to reading the comic strips when they are asked to report their favorite recreations. This is not invariably the case since occasionally "going to the movies" or "listening to the radio" is given a higher rank.* However, reading the comics surely represents a consuming and satisfying leisure pursuit of American boys and girls.

Some constructive efforts have recently concentrated on linking this interest with worthwhile educational endeavor. The makers of *My Weekly Reader* have attempted to provide a series of comic strips which can be readily associated with commendable activities in science or the social studies. Although the writer knows of no experimental evidence which demonstrates the value of this approach, he is convinced that the venture promises to be beneficial in capitalizing on a consuming interest of boys and girls. A still more direct approach has been fostered by the editors of *Parents' Magazine* who published an improved monthly magazine *True Comics*. So successful has this venture proved from a financial standpoint that this group has already provided other magazines ("Real Heroes," "Calling All Girls") for boys and girls.

All these efforts are based upon the recognition of a very real problem. To proceed most intelligently it seems necessary to have certain types of information which are at present unavailable. This paper represents an

* Lehman, H. C., and Witty, P. *The Psychology of Play Activities*. A. S. Barnes and Co., 1927.

** Hill, G. E., and Trent, M. E. "Children's Interests in Comic Strips." *Journal of Educational Research*, XXXIV, No. 1 (Sept. 1940), pp. 30–36.

* Eisenberg, A. L. *Children and Radio Programs*. New York: Columbia University Press, 1936. Cf. Witty P., and Kopel, D. *Reading and the Educative Process*. Boston: Ginn and Co., 1939. Cf. also Jersild, A., *Child Psychology* (Revised), New York: Prentice-Hall, 1940.

attempt to fill some gaps in the meager knowledge about the extent and nature of reading the comics in grades IV, V, and VI.

A questionnaire was devised and administered to 2500 pupils in grades IV, V, and VI in eight schools in Evanston, Chicago, and Milwaukee. In several schools in Evanston, Illinois, the questionnaires were employed as a basis for an interview, and additional data concerning reading habits, mental ability, educational attainment, nervous stability, and recreation were secured. In this paper, data will be presented for the children in the Evanston schools only.

During the interview, each child was asked to check from a list the comic books or magazines which he read, and to indicate whether he read them regularly, often, or occasionally. Similarly, he was given a list including the names of comic strips; these were discussed with him, and he was asked to mark the titles he read and to indicate how frequently he read them. He was then requested to indicate his favorites among the comic magazines and comic strips, and to add any titles not listed. Finally, additional information was secured concerning each child's inclination to make his own comic strips. Samples of the child's own endeavor were collected.

Number of Children Studied

In this paper, data will be presented for 334 boys and girls in the Dewey and Lincolnwood schools of Evanston. Table I shows the numbers of boys and girls in grades IV, V, and VI.

Table I

Number of Pupils

Grade	Number of Pupils		
	Girls	Boys	Both Sexes
IV	49	36	85
V	61	58	119
VI	61	69	130
Total	171	163	334

Results

Table II shows the average number of magazines or comic books read by the boys and the girls. It will be noted that the average for the boys is somewhat higher than that for the girls. The average for the entire group is about 13—a fact which attests to the popularity of comic magazines.

Table II

Mean Number of Comic Books or Magazines Read

Grade	Mean Number Read		
	Girls	Boys	Both Sexes
IV	11.08	14.97	12.73
V	11.03	15.57	13.24
VI	12.29	13.07	12.71
All grades	11.49	14.47	12.94

Many of these periodicals are published monthly. In order to ascertain something more about their reading habits, the boys and girls were asked to indicate how frequently they read these magazines. Table III presents data for this phase of the study. Boys appear to examine these magazines somewhat more frequently and with greater regularity than the girls.

Table III

Mean Frequency of Reading of Comic Magazines

Grade	Mean No. Read Regularly			Mean No. Read Often			Mean No. Read Some		
	Girls	Boys	Both	Girls	Boys	Both	Girls	Boys	Both
IV	2.02	4.19	2.94	3.26	4.11	3.62	5.79	6.27	6.00
V	1.32	4.74	2.99	2.11	4.07	3.06	6.51	6.91	6.70
VI	1.61	3.43	2.57	2.72	3.72	3.25	7.59	5.93	6.71
All grades	1.63	4.07	2.82	2.66	3.93	3.28	6.69	6.36	6.53

Thus, the boys read an average of four magazines *regularly* and about four others *often*. Girls, on the other hand, report *regular* examination of only one and one-half magazines; the average number which they peruse *often* is about two and one-half.

The most commonly read comics magazines are listed in Table IV. It is of interest that *Superman* and *Batman* attained first and second ranks for both boys and girls in all grades. Although there was a considerable degree of similarity in the ranks assigned the favorites in all grades, a few discrepancies occurred. *Action* was much more popular in grade VI than in grade IV. But the similarity is the conspicuous feature of these data as Table V shows. Neither in amount of reading

nor in the type of comic magazine preferred did there seem to be marked differences between grade IV, V, and VI. Moreover, the choices of the boys and the girls were similar, although the boys were attracted more frequently than the girls to comic books.

TABLE IV

RANKING GIVEN BY BOYS AND GIRLS TO THE TWELVE COMIC MAGAZINES GIVEN HIGHEST RANKING BY BOTH SEXES

Name of Magazine	Rank		
	Both Sexes	Girls	Boys
Superman	1.0	1.0	2.0
Batman	2.0	2.0	1.0
Famous Funnies	3.0	3.0	6.0
Action	4.0	8.0	3.0
Flash	5.0	7.0	4.0
Detective	6.0	12.0	5.0
Ace	7.0	6.0	8.5
Jungle	8.0	5.0	10.0
*Tip Top	9.0	4.0	13.0
True Comics	10.0	16.0	8.5
Donald Duck	11.0	9.5	11.5
Planet	12.0	9.5	13.5

* Not listed in questionnaire. Added by children in spaces provided for others.

TABLE V

FAVORITE COMIC MAGAZINES BY GRADES

(In Order of Preference)

Grade IV	Grade V	Grade VI
Superman	Superman	Batman
Batman	Batman	Superman
Flash	Famous Funnies	Famous Funnies
Famous Funnies	Flash	Action
Jungle	Ace	Tip Top
True Comics	Walt Disney	Detective
Donald Duck	Donald Duck	Flash
Tip Top	Wierd	Jungle
Magic	War	Ace
Detective	Super	Planet
Action	Detective	True Comics
Green Mask	Action	Donald Duck

The boys and girls reported that they read about the same number of comic strips. (Cf. Table VI.)

TABLE VI

MEAN NUMBER OF COMIC STRIPS READ

Grade	Mean Number Read		
	Girls	Boys	Both Sexes
IV	20.87	22.77	21.68
V	25.65	26.09	25.87
VI	26.47	23.79	25.06
Mean for all grades	24.57	24.38	24.48

The average number of comic strips which the boys and girls read *regularly* is only about half the total number which they examine. In addition to an average of thirteen comic strips read *regularly*, four or five additional ones are read *often* by these boys and girls, and six or more are *occasionally* read.

The data in Table VII attest unequivocally to the popularity of the comic strips among boys and girls in grades IV, V, and VI.

In tables VIII and IX, the twelve most popular comic strips are listed. One discovers that the three favorites occupy similar ranks in grades IV, V and VI. Discrepancies occur in the case of certain other strips, but, on the whole, there is a marked degree of correspondence in the favorites for the various grades. Although the sex differences are somewhat similar to those reported in other studies, they are less pronounced in the children's preferences for the comics than in other manifestations of reading interest.*

TABLE VII

MEAN FREQUENCY OF READING OF COMIC STRIPS

Grade	Mean No. Read Regularly			Mean No. Read Often			Mean No. Read Some		
	Girls	Boys	Both	Girls	Boys	Both	Girls	Boys	Both
IV	11.14	15.33	12.92	4.67	3.22	4.06	5.61	4.22	5.02
V	13.31	14.02	13.65	4.05	5.17	4.59	8.31	6.79	7.57
VI	12.85	13.84	13.18	5.44	4.06	4.71	7.59	6.13	6.82
	12.52	14.08	13.28	4.72	4.27	4.50	7.28	5.94	6.33

The children were asked to indicate whether they liked to make comics and to submit samples of their work. Table X shows that approximately 60 per cent of these boys and girls find pleasure in making comic strips. The variety of subject-matter covered and the generally high quality of the children's own products suggest important leads for teachers to follow in turning this interest in desirable channels.

* See Witty, P. A., and Kopel, D. *Reading and the Educative Process*. Boston: Ginn and Co., 1939.

TABLE VIII

RANKING GIVEN BY GIRLS AND BOYS TO THE TWELVE COMIC STRIPS GIVEN HIGHEST RANKING BY BOTH SEXES

Name of Comic Strip	Rank		
	Both Sexes	Girls	Boys
Dick Tracy	1.0	1.0	1.0
Smiling Jack	2.0	2.0	2.0
Blondie	3.0	3.0	4.0
Donald Duck	4.0	9.0	3.0
Captain and the Kids	5.0	9.0	5.0
*Brenda Starr	6.0	4.0	39.5
*Nancy	7.0	5.0	9.0
Terry and the Pirates	8.5	11.0	7.0
Mickey Mouse	8.5	16.0	6.0
Henry	10.0	9.0	9.0
Superman	11.0	13.0	9.0
Li'l Abner	12.0	13.0	13.5

* Not listed in questionnaire. Added by children in space provided for others not listed.

about 13. Of these an average of three magazines were read *regularly;* and three more were read *often.* Boys were inclined to show greater fondness than girls for comic magazines. The most frequently mentioned magazines were *Superman, Batman,* and *Famous Funnies;* issues of these were favorites in all grades.

Another indication of the interest in comics was found in the responses of the boys and girls to the comic strip. The average number of comic strips reported to have been read by the boys and the girls was about 25. Thirteen of these were *regularly,* and an additional 4 or 5 were *often* read. Occupying the first three ranks of favor in grades IV, V, and VI were *Dick Tracy, Smiling Jack,* and *Blondie.*

TABLE IX

FAVORITE COMIC STRIPS BY GRADES

(In Order of Preference)

Grade IV	Grade V	Grade VI
Dick Tracy	Dick Tracy	Dick Tracy
Blondie	Smiling Jack	Smiling Jack
Smiling Jack	Blondie	Blondie
Nancy	Donald Duck	Donald Duck
Gasoline Alley	Captain and the Kids	Brenda Starr
Donald Duck	Nancy	Terry and the Pirates
Tarzan	Brenda Starr	Captain and the Kids
Henry	Gasoline Alley	Mickey Mouse
Superman	Henry	Tiny Tim
Smitty	Mickey Mouse	Li'l Abner
Lone Ranger	Winnie Winkle	Henry
Captain and the Kids	Superman	Smitty

TABLE X

NUMBER AND PERCENTAGE OF PUPILS WHO LIKE TO MAKE ORIGINAL COMICS

Grade	Girls		Boys		Both Sexes	
	Number	Percentage	Number	Percentage	Number	Percentage
IV	30	61%	27	75%	57	67%
V	32	52%	29	50%	61	51%
VI	38	62%	39	56%	77	59%
All grades	100	58%	95	58%	195	59%
Mean	33.33	58+%	30.66	60+%	65	59%

SUMMARY

This study of 334 boys and girls in grades IV, V, and VI reveals that reading the comics is one of the pupils' favored leisure activities. In fact, it appears to be the most popular of *all* reading pursuits.

Interest in the comics appears in many forms. Reading comic books or magazines is a popular pastime. The average number of magazines reported to have been read was about 13. Of these an average of three magazines were read *regularly;* and three more were read *often.* Boys were inclined to show greater fondness than girls for comic magazines. The most frequently mentioned magazines were *Superman, Batman,* and *Famous Funnies;* issues of these were favorites in all grades.

Another indication of the interest in comics was found in the responses of the boys and girls to the comic strip. The average number of comic strips reported to have been read by the boys and the girls was about 25. Thirteen of these were *regularly,* and an additional 4 or 5 were *often* read. Occupying the first three ranks of favor in grades IV, V, and VI were *Dick Tracy, Smiling Jack,* and *Blondie.*

A final indication of the influence the comics exercise on boys and girls was revealed by their tendency to make original comics. This tendency was clearly shown in the fourth grade, where about two-thirds of the youngsters reported participation in this activity. In grades V and VI, the interest persisted—about 60 per cent of the children in these grades indicated that they found pleasure in making their own comics.

All these activities seem to satisfy the middle grade child's desire and need for experiences that are adventurous and exciting. These elements seem much more important than the item of humor, since many widely read comics are devoid of humorous situations or incidents.

Children's interests in the comics are similar to those found: in other forms of reading, in the cinema, and in the radio. The appeal of all three media is identical.

These data reveal a need for a realistic approach on the part of teachers and parents in dealing with the comics. The teacher faces a problem similar to that presented by the radio and by the cinema. In meeting it, it is essential that the teacher make a rather complete appraisal of each child's interests and his many-sided nature, since it is only from an examination of the child's total pattern that cues may be obtained for directing him. In some children, an allotment of time to the comics which seems excessive may be viewed otherwise in terms of a well-chosen and judicious balance of worthy interests and pursuits. In other children, a relatively small amount of interest may seem unwarranted in terms of an otherwise impoverished recreational life. Techniques such as the interest-inventory, the anecdotal record, or the child's own diary offer the teacher avenues through which the child's total behavior pattern may be studied.

In the middle grades, teachers might well proceed a step further, and attempt to surround children with a variety of good literary sources which are rich in the elements of action, surprise, adventure, and excitement. Such material is readily available at the present time: in publications such as the *Disney Readers*, the *Story Parade Adventure Books*, and the *New World Neighbors*.* In the first two sets of books, an effective antidote for the current craze for comic magazines is found, and in the third series, one may find appropriate material for the cultivation of permanent interests in good books.

In another paper, the writer will present case studies which reveal how this guidance has been effectively administered. In this paper he has sought merely to describe the nature and extent of interest in the comics among middle grade children.

* *The Disney Readers* and the *New World Neighbors* are published by D. C. Heath and Co.; *Story Parade Adventure Books*, by Grosset–Dunlap.

READING THE COMICS—A COMPARATIVE STUDY

PAUL WITTY
Northwestern University

In another manuscript[1] the writer presented data concerning children's interest in the comics. Three hundred thirty-four boys and girls in grades IV, V, and VI of the Evanston, Illinois schools were interviewed. The investigation disclosed a genuine interest in comic books, magazines, and strips in all grades. The average number of magazines read was about thirteen; three were read *regularly;* an additional three were read often, and six were read *occasionally*. Reading the comic strips was also a popular pursuit; an average of twenty-five comic strips were read. Of these, thirteen were read *regularly* and four were examined *often*. Finally, about sixty per cent of the pupils reported that they enjoyed making their own comics. These data attest unequivocally to a general interest in comics among middle-grade children, and they suggest the validity of the statement that the comics represent children's major reading interest.

The purpose of this investigation is to study the interests of children selected at random from 2500 public school youngsters on the same points as those covered by the preceding investigation. Children in grades IV, V, and VI of eight schools in Chicago and vicinity were interviewed. Samples of one hundred each were made at random by including the reports of every twentieth child in the first sample, and proportionate numbers in other samples. Eliminations and additions were made until the samples were constituted in the manner shown in Table I. The samples showed remarkable agreement. Hence, it seemed that the method provided a reliable picture of the entire group of 2500 boys and girls.

[1] See the preceding paper in this number of the Journal. Ed.

TABLE I

NUMBER OF PUPILS IN RANDOM SAMPLING

Grade	Number of Schools*	Number of Pupils		
		Girls	Boys	Both Sexes
IV	3	18	15	33
V	4	17	16	33
VI	5	16	18	34
All grades	7	51	49	100

* Names of schools: Seward, Hammond, Falconer, Motley, Haven, Anderson, Field.
Names of towns: Evanston, Chicago, Illinois; Milwaukee, Wisconsin.

TABLE II

MEAN NUMBER OF COMIC MAGAZINES READ

(Random Sampling)

Grade	Mean Number Read		
	Girls	Boys	Both Sexes
IV	16.33	15.80	16.09
V	14.94	16.87	15.87
VI	11.81	16.89	14.50
All grades	14.45	16.55	15.48

Table II shows the number of comic magazines read by these boys and girls. It will be noted that the average number is slightly above 15, and that the boys read a considerably larger number than the girls. The Evanston group read a somewhat smaller number (13), and differences between the sexes were similar to those in the random sampling.

Table III shows that an average of four magazines were read *regularly;* four and one-half, *often;* and about seven *occasionally*.

TABLE III

MEAN FREQUENCY OF READING OF COMIC MAGAZINES

(Random Sampling)

Grade	No. Read Regularly			No. Read Often			No. Read Sometimes		
	Girls	Boys	Both	Girls	Boys	Both	Girls	Boys	Both
IV	6.05	4.06	5.12	4.44	4.20	4.33	5.83	7.53	6.61
V	2.11	3.62	2.84	5.82	4.75	5.30	6.88	8.50	7.66
VI	3.06	4.11	3.62	3.12	5.44	4.06	5.62	7.33	6.53
All grades	3.80	3.93	3.87	4.49	4.84	4.66	6.12	7.77	6.93

This study is similar to others in showing that boys read comic magazines somewhat more frequently than girls. But the differences are neither large nor consistent from grade to grade.

TABLE IV

RANKINGS GIVEN BY GIRLS AND BOYS TO THE THIRTEEN COMIC MAGAZINES GIVEN HIGHEST RANKING BY BOTH SEXES

(Random Sampling)

Name of Magazine	Rank		
	Both Sexes	Girls	Boys
Batman	1.0	2.0	1.0
Superman	2.0	1.0	2.0
Jungle	3.0	3.0	9.0
Famous Funnies	4.0	4.0	9.0
Flash	6.0	10.5	4.0
Ace	6.0	5.0	11.0
Magic	6.0	6.0	6.5
Action	8.0	13.0	4.0
Detective	9.0	*	4.0
True	10.0	13.0	9.0
Green Mask	12.0	8.0	*
Shadow	12.0	8.0	*
Wings	12.0	*	6.5

* Listed by one person.

The favorite comic magazines are displayed in Table IV. In general, there is considerable similarity in the choices of the boys and the girls. The titles of favorite magazines are very similar to those in the Evanston group. Yet in the random sampling, there is a somewhat larger range with a higher frequency of magazines which appear only once.

Magazines favored in the different grades are listed in order in Table V. The similarity in the choices of these grades is the essential feature of this table. Close inspection shows that the magazines *Jungle*, *Magic*, and *Famous Funnies* occupy somewhat different positions in the three grades. Despite this discrepancy, we must conclude that the most popular magazines have a rather general and similar appeal in every grade.

TABLE VI

MEAN NUMBER COMIC STRIPS READ

(Random Sampling)

Grade	Mean Number Read		
	Girls	Boys	Both Sexes
IV	27.67	25.73	26.79
V	26.05	24.25	25.18
VI	27.00	26.00	26.47
All grades	26.92	25.34	26.15

Table VI presents the average number of comic strips which the children read. It will be observed that the average number both for the boys and the girls is about 26. Moreover, this number is found in the fourth grade and persists through the sixth grade. Thus the intense interest which is evident in the fourth grade is not lessened during the ensuing two grades.

Table VII shows that these girls read the comic strips more frequently than the boys. This somewhat contradictory finding probably reflects the fact that sex differences are not as pronounced or consistent in reading the comics as in other reading pursuits. In several of our tabulations, boys have been found to read the comics more frequently than the girls. But, in all these tabulations, the differences between the sexes are neither

TABLE V

FAVORITE COMIC MAGAZINES BY GRADES

(Random Sampling)

Grade IV		Grade V		Grade VI	
Name	Rank	Name	Rank	Name	Rank
Batman	1.0	Batman	1.0	Batman	1.0
Superman	2.0	Superman	2.0	Superman	2.0
Famous Funnies	3.5	Jungle	3.5	Flash	3.0
Magic	3.5	Ace	3.5	Action	4.5
Jungle	5.0	Action	5.5	Famous Funnies	4.5
Flash	6.5	Detective	5.5	Jungle	6.5
Shadow	6.5	Crack	7.5	Ace	6.5
Green Mask	10.5	Adventure	7.5	Detective	10.5
Planet	10.5	Flash	10.5	Mystery	10.5
Ace	10.5	Famous Funnies	10.5	True	10.5
True	10.5	Fantastic	10.5	Magic	10.5
Wings	10.5			Tip Top	10.5
Blue Beetle	10.5			Wings	10.5

TABLE VII

MEAN FREQUENCY OF READING OF COMIC STRIPS

(Random Sampling)

Grade	Number Read Regularly			No. Read Often			No. Read Sometimes		
	Girls	Boys	Both	Girls	Boys	Both	Girls	Boys	Both
IV	15.83	13.07	14.57	6.77	2.73	4.93	3.66	9.93	6.51
V	16.06	13.00	14.57	4.00	4.18	4.09	6.00	6.12	6.36
VI	14.87	14.61	14.73	4.56	6.44	5.56	7.31	4.77	5.97
All grades	15.60	13.61	14.63	5.16	4.57	4.87	5.59	6.79	6.28

consistent nor highly significant. The only justifiable conclusion seems to be that boys and girls are generally attracted to the comics. The attraction occurs in the early grades and continues with great consistency throughout the middle grades.

Table VIII presents the most popular comic strips. It may be seen that the preferred strips are popular in grades IV, V, and VI. The differences between the sexes are relatively insignificant. It is apparent that certain strips have strong appeal in all three grades. The fact that reading the comics has general rather than specific appeal may be gathered from observing that grade differences are small, sex differences are insignificant, and differences in locality

TABLE VIII

RANKINGS GIVEN BY GRADES IV, V AND VI TO THE COMIC STRIPS GIVEN HIGHEST RANKING BY ALL GRADES

(Random Sampling)

Name of Comic Strip	Rank			
	All Grades	Grade IV	Grade V	Grade VI
Blondie	1.0	1.0	2.5	1.0
Dick Tracy	2.0	13.0	1.0	2.0
Smiling Jack	3.0	2.0	2.5	5.0
Lone Ranger	4.0	5.0	8.0	3.0
Donald Duck	5.0	3.0	13.5	5.0
Mickey Mouse	6.0	13.0	5.0	5.0
Brenda Starr	7.0	7.5	13.5	13.0
Henry	11.0	5.0	13.5	—***
Toots and Casper	11.0	—**	13.5	8.5
Dixie Dugan	11.0	13.0	13.5	13.0
Captain and the Kids	11.0	—**	8.0	13.0
Winnie Winkle	11.0	—**	4.0	—***
Terry and the Pirates	11.0	—***	8.0	8.5
*Scarlet O'Neil	11.0	13.0	8.0	—**

* Not listed on questionnaire. Added by children in space provided for others.
** Listed as a favorite by only one person.
*** Not listed as a favorite by any child.

appear to influence the amount and nature of this type of reading to a small degree only. The latter may be seen by noting the similarity in the nature and extent of reading the comics between the random sample and the Evanston group.

This similarity is vividly shown by comparisons of Tables VIII and IX, which present the favorite comic strips of the two groups.

The general interest of boys and girls in the comics is revealed also by their inclination to make comics. In fact, exactly half of these boys and girls report a fondness for making their own comics. Again, a relatively greater interest of the boys is revealed; 55 per cent of the boys and only 45 per cent of the girls indicated that they enjoyed this activity.

It is of interest that Evanston boys and girls read a somewhat smaller number of comic magazines and comic strips than the unselected group; however, they make their own comics more frequently and turn to the comics early in the grades and persist in reading them.

TABLE IX

RANKINGS GIVEN BY GRADES IV, V, AND VI TO THE COMIC STRIPS GIVEN HIGHEST RANKINGS BY ALL GRADES (EVANSTON)

Name of Comic Strip	Rank			
	All Grades	Grade IV	Grade V	Grade VI
Dick Tracy	1.0	1.0	1.0	1.0
Smiling Jack	2.0	3.0	2.0	2.0
Blondie	3.0	2.0	3.0	3.0
Donald Duck	4.0	6.0	4.0	4.0
Captain and the Kids	5.0	11.5	5.0	7.0
*Brenda Starr	6.0	20.0	7.0	5.0
*Nancy	7.0	4.0	6.0	12.5
Terry and the Pirates	8.5	15.5	16.5	6.0
Mickey Mouse	8.5	20.0	9.0	8.0
Henry	10.0	8.0	9.0	11.0
Superman	11.0	8.0	12.5	17.0
Li'l Abner	12.0	15.5	14.5	10.0

* Not listed in questionnaire. Added by children in space provided for others not listed.

TABLE X

NUMBER AND PERCENTAGE OF CHILDREN WHO LIKE TO MAKE ORIGINAL COMICS

(Random Sampling)

Grade	Girls		Boys		Both Sexes	
	Number	Percentage	Number	Percentage	Number	Percentage
IV	8	44%	8	53%	16	48%
V	8	47%	11	68%	19	57%
VI	7	44%	8	44%	15	44%
All grades	23	45%	27	55%	50	50%
Mean	7.66	45%	9.00	55%	16.66	49⅔%

During the past two years, a number of dogmatic statements have been made concerning the baneful influence of reading the comics upon the development of boys and girls. The practice is sometimes condemned with such finality that parents and teachers are inclined to believe that reading the comics should be discouraged or actually forbidden. In an effort to study the validity of these strictures, the ten per cent of the pupils who read comics most in one school were compared with the ten per cent who read them least. The average IQ. of the two groups was in the interval 105–110 (most 107; least 105).

The boys and girls in the two groups received almost the same average marks and were considered by their teachers to be about equally well-adjusted and effective in social relationships. When the general reading of the children who turn to the comics most frequently was compared with that of the group reading the comics least often, it was found that the patterns were quite similar.

Case studies also revealed that some children who read the comics very often were following reading programs which were varied, rich, and generally commendable. Thus, one sixth grade boy, who reported that he read 28 comic magazines *occasionally* and more than 20 strips *regularly*, listed 30 books which he had read voluntarily during the few months preceding the investigation. In this list were four new volumes dealing with South America, two *Story Parade Adventure Books*, several good biographies, and fair assortment of exciting stories such as *Cowboy in the Making* by Will James. In addition, he had read recently a number of magazines and a considerable amount of poetry including *Peacock Pie and Other Poems* by De la Mare; *Rainbow in the Sky* and others by Lindsay, and several poems by Sandburg.

On the other hand, a boy and a girl who read very few comics and two children who read a very large number were generally inadequate. Although their intelligence quotients were average, their reading patterns were distinctly impoverished.

In dealing with the pupils who were preoccupied with the comics, the cooperation of the home was sought in an effort to make other types of reading matter equally satisfying and enjoyable. The first step involved ascertaining rather fully the varied interests and leisure pursuits of these children. A number of books, magazines, and pamphlets were then made available. These were selected to satisfy or extend the children's interests. Since these children had become habituated to exciting, adventurous books, radio programs, and movies, an effort was made to capitalize on these interests and to introduce them slowly and progressively to good literature. Among the materials that were advantageously used in transitional stages were the *Disney books*, the *Story Parade Adventure Series*, and twenty volumes on aviation and sports. Reading the Disney books constituted the first new reading experience for these children. Another phase of the program included the cultivation of other inchoate interests. One child found her interest in music happily satisfied by hearing *Peter and the Wolf* and other records which caused her to appreciate new possibilities in records and in the radio; and books about music and musicians provided an easy introduction to more worth while reading.

One boy who read very few comics, displayed great originality and considerable skill in making his own comics. This interest was directed into closely related types of artistic endeavor, and it was associated with books which contained information about these new pursuits. With every pupil, the procedure was similar. First, there was a

careful study of the child's interests and needs; and second, there was an effort to find worthy reading material to satisfy every child's interests. In some cases, it was necessary to create new interests through an enrichment and extension of experience. But, in every instance, the results were gratifying.

Summary

Reading the comics occupies an important place in the list of children's favorite leisure pursuits. A random sampling procedure indicated that, among 2500 children in grades IV, V, and VI, the interest in comics was general and consistent from grade to grade. The average number of magazines read was about 15 in each of the grades; four of these were read *regularly;* and four and one-half, *often*. Comic strips also were generally read; twenty-six was the average number reported to have been read. Of these, about 15 were read regularly; and 5 often. Sex differences and grade differences were not pronounced. Thus, the data reveal a *general* interest in reading the comics. The interest is well-defined in the fourth grade, and is maintained with about the same intensity in the fifth and sixth grades. Moreover, the same comic magazines and strips are popular in all groups.

This interest appears in several forms. In one of these, making their own comics, the boys show a more marked inclination than the girls. About half the children make their own comics.

These data secured from the random sample are similar to those obtained for a more highly selected group from two schools in Evanston, Illinois. The differences between these groups are smaller than one would anticipate from comparisons of other manifestations of reading interest. They support the body of data showing that at the present time, reading the comics represents a general interest which in grades IV, V, and VI is relatively uninfluenced by differences in age or grade, sex, or locality.

Strictures are often made concerning the dire effects of reading of the comics. The data in this paper do not afford a basis for these diatribes. In one school, two groups of children were studied intensely—those who read comics extensively and those who seldom read them. Reading patterns of the two groups disclosed little difference in amount or nature of reading. Case studies indicated that one could not generalize concerning the needs of children from a single tendency such as maximum or minimum amount of reading in this area. Instead, the problem of creating desirable reading habits can be met only by appraising each child's *entire* reading pattern. For some children, an amount of reading in the field of the comics may seem excessive. However, when it is viewed in terms of an otherwise well-rounded and balanced program of reading, it can scarcely be considered harmful.

Similarly, a small amount of reading of comics may be associated with a desirable or with an undesirable *total* pattern. It seems that a child's tendency to read the comics should not be used independently as a basis for appraising his needs. Guidance should be based upon a more extensive and far-reaching attempt to understand each child's needs for reading and his attempts to satisfy them through different avenues. Finally, since reading the comics is actuated by the same motives which are satisfied by the radio and the movie the problem is one which apparently can not be dealt with singly. Nevertheless, appraisal and guidance of this fundamental reading habit are responsibilities of every teacher.

WORDS AND THE COMICS

ROBERT L. THORNDIKE
Teachers College, Columbia University

Within the last decade, a new form of reading material has taken a dominant place in the experience of children. This is the "comic" book, more than 100 of which are published month by month, some with circulations running into the millions. The comic books have not had an entirely friendly reception among parents and educators, and certainly they do not represent the material which most educated adults would choose for their child to read, but they have been warmly received by children. Since they have such a strong appeal to children, and are so widely read by them, it seems worth while to examine somewhat more analytically the reading experiences which they provide.

There are many angles from which comic books might be studied. One approach would center around themes and plot patterns. Another would focus on characterizations. A third on social attitudes expressed or implied. All of these would be valuable. However, the interest in this inquiry is upon the range and difficulty of the vocabulary which is encountered in these books. How much reading do comic books provide? What, approximately is their level of difficulty? How much, and what type of vocabulary experience does the child encounter who reads them?

It would be manifestly impossible to make an analytical study of all the hundred-odd magazines which now flood the newsstands. This study was therefore limited to copies of four magazines whose circulation is among the highest, all products of one publisher. These were *Superman, Batman, Action Comics,* and *Detective Comics.* These four, together with a number of lesser magazines, are all published by *D. C. Publications.* Specifically, a detailed vocabulary count was made of *Superman No. 9, Superman No. 11, Batman No. 6,* and *Detective Comics No. 53.* Short selections were analysed for reading difficulty from other issues of these magazines.

In view of the type of sample of magazines used, we must hesitate to generalize our results too widely to the host of other magazines, some rather fly-by-night publications. However, the consistency of the results for the four issues studied suggests that they provide a fairly adequate picture of this particular group of well-established and widely circulated magazines.

AMOUNT OF READING

How much reading experience does the child get when he reads one of these comic magazines? A count of the number of words in the text shows it to cluster around 10,000. To the nearest hundred, we found the following wordage:

Superman #9	10,300
Superman #11	9,300
Batman #6	10,600
Detective Comics #53	9,800

Many readers will probably be surprised that the amount is as large as this, thinking of the comic books as essentially picture books. The child who reads a comic book once a month through the school year (and this represents a very moderate dosage), gets about as much wordage of reading as he gets from even the new fourth or fifth grade reader. In view of the need of the upper elementary school and junior high school pupil for a large volume of reading and vocabulary-building experience, this source should not be neglected.

VOCABULARY CONTENT

A complete word count was made, for these four comic books, of the occurrences of all words outside the first thousand in the Thorndike list of 20,000.* No tabulation was made of words in the first 1,000: (1) because of the large amount of labor involved in counting these very common words and (2) because these commonest words are of little importance in increasing either reading or comprehension vocabulary. In Table I is presented a summary of the vocabulary content of the four books analysed. This table shows the number of different words from each level of commonness in the Thorndike

* Thorndike, E. L., Teacher's Word Book of 20,000 Words; Bureau of Publications, Teachers College, Columbia University, N. Y.; 1932.

TABLE I

NUMBER OF DIFFERENT WORDS FROM VARIOUS THOUSANDS OF THORNDIKE LIST IN FOUR COMIC BOOKS

(Number of Different Words)

Thousands of Thorndike List	Superman #9	Superman #11	Batman #6	Detective Comics #53	Combined
2	296	238	242	231	580
3	174	147	149	134	389
4	126	116	83	80	278
5	88	83	64	65	210
6	66	63	59	45	168
7	52	51	36	36	133
8	41	41	36	18	114
9	35	35	22	22	89
10	30	25	19	20	70
11–15	88	76	56	77	234
16–20	41	34	33	52	129
Not in Thorndike List	185	154	198	196	649
Total	1222	1063	997	971	3043

list for each book, and for the four books put together. In the composite list, words which occurred in more than one book are, of course, counted only once.

As we examine Table I, we see that each book contains about 1000 different words other than those in the commonest thousand. Of these, about 500 fall in thousands 2–5, 200 in thousands 6–10, 100 in thousands 11–20, and 200 outside the 20,000 in the Thorndike list. The four books together give about 3000 different words outside the first thousand. The child who reads these four books would encounter about 1000 different words falling in the 5th to 20th thousands of the Thorndike list—i. e., uncommon words in the reading of the elementary school child, but words which he should probably be encouraged to learn.

That additional new words would be added by additional books is suggested by Table II. It is seen in this table that each new book adds a group of additional words to the vocabulary already encountered, and that new words are still being encountered in substantial numbers in the last book of the four. The vocabulary of the comics is in no sense exhausted by these four books, and additional issues would certainly continue to add a substantial number of new words.

The 3000 different words outside the first thousand of the Thorndike list represent a total of about 12,000 occurrences. On the average, then, each word occurs four times. The frequencies are very unevenly distributed over the different words, and most of the words outside the first two or three thousands will, of course, need a good deal of additional practice beyond what is provided in this small reading sample.

The type of vocabulary encountered is best illustrated by samples. The lists below give the following samples: (1) the first (alphabetically) 20 words in thousands 2–5, (2) the first 20 words in thousands 6–20, (3) the first 20 words not included in the Thorndike list.

Words From Thousands 2–5	Words From Thousands 6–20	Words Not In Thorndike List
abandon	abreast	abduction
ability	abrupt	adios
aboard	absorb	adrenalin
accept	absorbing	after-dinner
accident	accidental	aide
accompany	accomplice	aileron
accomplish	accordance	aircraft
according	accurate	airplane
accuse	accusation	airport
acknowledge	ace	Al
acquaint	acquit	Alaskan
action	acrid	allergic
active	activity	an'
address	actress	Angelica
adjoin	actually	anti-aircraft
admire	addict	anymore
admit	adjourn	applauder
advance	adjust	arrumph
adventure	administer	Atkins
advice	affected	atop

Of special interest are the words not included in the Thorndike list. A special analysis, admittedly somewhat subjective in char-

TABLE II

NEW WORDS IN EACH ADDITIONAL COMIC BOOK

(Words First Encountered in)

Thousand of Thorndike List	Superman #9	Superman #11	Batman #6	Detective Comics #53
2– 5	684	304	265	204
6–10	224	145	114	91
11–15	88	55	41	50
16–20	41	24	29	35
Not in List	185	138	173	153
Total	1222	666	622	533

acter was made of these words. By inspection it was found that certain classes of words recurred in this group and categories were set up for analysis. The categories with examples, and the approximate proportion of each are shown below.

Standard Words (abduction, atop, awash)	23%
Words of recent currency (aileron, airport, autogyro)	5%
Home-made compounds (beetle-browed, bludgeon-blows)	21%
Contractions (an', b'low, bein')	6%
Proper Names (Angelica, Atkins, Balconia)	21%
"Respectable" slang (adios, brainstorm, buddy)	12%
"Vulgar" slang (awk, betcha. conk)	10%
Miscellaneous	2%

Of these 649 words, a fair number appear of positive value and worth encountering, many appear essentially neutral—neither of value nor of harm to know, and only a minority appear to be possibly objectionable. However, the total amount of slang in the books is somewhat greater than the 22% of these non-listed words. This is because a number of standard words are used with slang meanings. A word count fails to reveal slang usages of such words as rod, stir, bunk, copper, joint, etc. But even when this is allowed for, one hardly gets the feeling that the language of these particular magazines was excessively slangy.

READING DIFFICULTY

An effort was made to estimate the reading difficulty of the text in these comic books. The usual formula for estimating grade level of difficulty may not be directly applicable to this material because of (1) the disconnected character of the text and (2) the aid given by the picture context. However it is interesting to see how the reading would rate, just as reading. Several 300-word samples have been selected from different comic books, and evaluated by means of the Lorge formula.* The grade level of difficulty of the selections studied was as follows.

Superman #11 p. 3 ff.	6.6
Superman #11 p. 23 ff.	5.7
Action Comics #42 p. 11 ff. (Superman)	7.4
Action Comics #41 p. 23 ff. (Black Pirate)	6.2
Action Comics #41 p. 36 ff. (Mr. America)	5.5
Batman #8 p. 4 ff.	5.9
Batman #8 p. 61 ff.	5.6
Detective Comics #53 p. 6 ff. (Batman)	5.2
Detective Comics #56 p. 21 ff. (Crimson Avenger)	6.0
Detective Comics #57 p. 42 ff. (Speed Saunders)	5.1

With one exception, these selections fall at fifth or sixth grade reading difficulty according to the formula. This is consistent with the group who read them, because the bulk of the readers fall within the age range of 9 to 15 years. They represent material of difficulty appropriate to the average 11- or 12-year-old, and to the slow child of somewhat greater age. One would infer that simply as reading, apart from reference to the pictures, these books represent a fairly substantial task for a number of their patrons. The difficulty is probably reduced somewhat, and the interest appeal increased greatly, by the picture context, so that the slow children especially are led to read with enjoyment material which would be difficult for them if presented as straight text. If this conclusion is justified, its significance for those working in remedial reading seems to be obvious.

In this report, attention has been paid only to the *words* of the comics, not to the *ideas;* to the structure, not to the content. This must not be construed as belittling the importance of other questions. Whether the comics provide exposure to a viciously distorted and unrealistic world, whether they merely provide a rather innocuous way of wasting children's time, or whether they provide a needed vicarious release for tensions and aggressions which are built up and unexpressed in the world of reality is a *vital* question. It is also a *moot* question, and not one which we are able to resolve here. It must be explored critically but sympathetically by educators, psychologists, and psychiatrists.

* Lorge, Irving. "Predicting Reading Difficulty of Selections for Children," *Elem. Eng. Rev.*, XVI, (1939) 229–33.

Conclusions

In summary, a vocabulary analysis of four complete comic books and an evaluation of the difficulty of selections from other issues revealed the following facts:

(1) Each book contained about 10,000 words of reading matter.

(2) Each book contained about 1000 different words other than those falling in the commonest thousand of the Thorndike word list.

(3) The four books together contained about 3000 different words other than those in the first thousand of the Thorndike list. As each book was added on, additional new words continued to be encountered.

(4) Although a number of slang words and slang meanings were encountered, the bulk of the vocabulary was standard English. Many hundreds of words were included which the child needs to encounter as he expands his reading vocabulary.

(5) The reading difficulty of the text material, as estimated by the Lorge formula, was at fifth and sixth grade level.

It appears, then, that the "comics", whatever their other vices or virtues, *do* provide a substantial amount of reading experience at about the level of difficulty appropriate for the upper elementary school or even junior high school child. In view of the apparent interest and appeal of this material for the child, the facts presented above suggest that this supplementary resource may have real value for the educator who is interested in working with the child as he is and in leading him on from his present status to higher and better things.

AN APPROACH TO THE MEASUREMENT OF INSECURITY

ELINOR PRICHARD and RALPH H. OJEMANN
Child Welfare Research Station
University of Iowa

The term "insecurity" and its correlative "desire for security" appear extensively in child development literature. Insecurity is listed frequently among the possible causes of many behavior problems. Careful analysis of the literature, however, reveals that there is little agreement as to the meaning of these terms and that methods of identification and measurement of insecurity are not well developed. Since it appears highly probable that desire for greater security plays an important part in the behavior of children, it would be most helpful if we could develop more precise methods of classifying children with respect to the strength or frequency with which this desire motivates their behavior. We need methods by which we can discriminate between relatively secure and the relatively insecure children.

The problem of measuring desire for security, as all measurement of factors underlying behavior, involves the interesting question of the relation of behavior as it can be observed to the factors that underlie or cause that behavior. It is clear that the way a child will express his "seeking for security" will depend in part upon what psychological resources he has at his command, i. e., what skills, ideas and energy he has available. A baby who cannot walk will not show a feeling of insecurity by walking away from the group. He will show it in other ways such as turning toward mother and crying. The behavior may take as many different forms as there are variations in psychological equipment.

The problem appears further complicated when we consider that in studying such a want as desire for security we are concerned with motive rather than with some item of equipment. In measuring knowledge or skill we are measuring essentially an item of resource or equipment. In making measurements of knowledge and skill we put the child in a situation in which we control motive by asking each child to do his best and the test is not considered a satisfactory one unless there are evidences that the child is putting forth his full effort. When measuring motive, however, we must use situations in which the motive is not controlled by the experimenter but by the subject since it is the subject's own hierarchy of wants that we wish to examine.

In developing methods for identifying the relative strength or position of the desire for security in the subject's hierarchy of motives we have then two problems: on the one hand, we must make the measurement as independent of variation in resources or equipment as we can. Secondly, we must use situations in which the hierarchy of wants of the subject is not directly controlled by the experimenter.

Another way of stating these problems is to think of the relation of behavior to the factors that underlie behavior in the form of an equation such as:

$$\text{Motive} + \text{equipment} + \text{immediate environmental setting} = \text{Behavior}$$

If we think of this equation as a dynamic expression, i. e., the plus signs are the plus signs of dynamics and not the simple additive signs of arithmetic, it seems that if we wish behavior to vary with variations in one member of the equation, the other two must be held constant or their effect must be uniform.

In the method of measurement developed in this paper an attempt was made to minimize the effect of variations in equipment and variations in the immediate environmental setting by such procedures as the following: Since preschool children served as subjects, only those activities were considered which required simple skills possessed by all children of the preschool level. Bidding for the attention of the adult either verbally or by other methods, withdrawing from a group, crying, revealing tenseness in facial expression, and clutching adult for help when facing a new situation are examples of activities requiring abilities common to all pre-

school subjects. A variety of patterns was observed to further minimize the effect of variation in resources.

To minimize the effect of variations in immediate environmental setting the environmental conditions under which the behavior occurred was specified. For example, one type of behavior evidenced by insecure children is withdrawal from the group. A child may withdraw from a group because he doesn't feel secure in it, but he may also withdraw to investigate some noise that he has just heard, to participate in some interesting activity that has distracted him, or for many other reasons. If withdrawal is to be used, it would seem desirable to eliminate instances of withdrawal from a group resulting from such distractions as leaving a group to secure play materials, to answer the teacher's call, to investigate interesting sounds, to manipulate attractive play materials that have come into view. It is not withdrawal from a group under any condition but withdrawal from a group under specified conditions that will serve as a useful item in the repertoire of indicators of insecurity.

If the above reasoning is correct, we should find behavior patterns appearing under specified conditions that differentiate secure and insecure children consistently. Can we isolate such patterns?

This study was carried out in a preschool environment. Two groups of children, one labeled "secure" by their respective teachers and the other labeled "insecure," were observed for a definite period of time. Records were made of their behavior and the results analyzed to determine which items reliably differentiate "secure" from "insecure" children. The details of the method follow.

Two preschool teachers, the head teacher and one other in each of two different groups at two different midwest institutions, were asked to select the four most insecure and the four most secure children in their respective groups. The teachers were requested to work independently, and were told that their judgment was the only criterion that would be considered. It was found that the agreement between the teachers was rather high. At one institution the agreement was complete for all the secure cases and in seventy-five per cent of the insecure cases. At the second institution there was complete agreement on the insecure children and in seventy-five per cent of the secure children. In all, thirty children were observed, fifteen insecure and fifteen secure. Fourteen were from one institution and sixteen from another.

Five ten-minute periods of observation were obtained for each child on five consecutive days that the child was present. An effort was made to include a variety of situations, both free play and routine. Children were observed individually and the child's behavior was recorded as a running account in shorthand. At this point no effort was made to select behavior patterns and the analyses were not begun until all observations were made.

To obtain an approximate indication of the reliability of the observers early in the study, seven ten-minute periods of observation were made with another observer on seven children. A group of nine patterns which seemed from preliminary observation and reading to occur frequently in insecure children were tabulated as to frequency from each observer's record. The number of recorded incidents of the nine behavior patterns for each subject as determined by each of the two observers were correlated. A coefficient of .94 was obtained.

The observations of the secure and insecure children were then analyzed to determine the behavior patterns that occurred frequently in either group. A total of six items seemed to stand out. These items were carefully described as to conditions under which they appeared, and the frequency tabulated for the "secure" and "insecure" group. The detailed descriptions of the items, the frequency of occurrence, and the value of "t" for the differences between the means of the total group are given in Table 1. The subjects ranged in age from two years, six months to five years, six months and could be conveniently classified into two age groups —those above four years and those below. The table shows the mean frequency for the "secure" children and "insecure" children in each age group, the mean frequency for all "secure" and "insecure" children and the "t" of the differences between total "secure" and "insecure." An inspection of the data in the last column indicates that five of the six patterns show a value for "t" beyond 2.763 required for the one per cent level of confidence for twenty-eight degrees of freedom. In interpreting the t-test in a study of this

TABLE I

MEAN FREQUENCY OF BEHAVIOR PATTERNS IN "SECURE" AND "INSECURE" GROUPS

Pattern	Mean Frequency in Younger Group 2 years, 6 months Insecure N=7	Mean Frequency in Younger Group 4 years, 0 months Secure N=7	Mean Frequency in Older Group 4 years, 0 months Insecure N=8	Mean Frequency in Older Group 5 years, 6 months Secure N=8	Mean Frequency in Total Group Insecure N=15	Mean Frequency in Total Group Secure N=15	t
Withdrawal—This is indicated by an evident preference for playing alone. A child who makes very few, if any, attempts to join a group. This does not apply to a child who withdraws from a group to follow some other interest or in response to observable stimuli.	23.71	4.42	23.87	4.87	23.80	4.66	7.79*
Non-acceptance—This applies to a child who wants to join the group and tries unsuccessfully to do so. Rejection may be physical or verbal. It does not apply to a child who watches a group at play but makes no overt attempt to join it.	5.85	1.28	7.62	2.50	6.08	1.93	3.92
Bidding for attention—This applies to a child who is constantly aware of an adult's presence, and who continually appeals to the adult for attention. This may be manifested in such ways as following adult, seeking unnecessary help, appealing for praise and approval, or any other conduct of a similar nature. It does not apply to a child who asks needed help, or who watches the teacher when she is talking to the group or showing them something.	29.14	11.85	24.37	12.12	26.60	12.00	5.46
Crying—Frequent crying in situations which do not ordinarily result in such behavior, such as crying when reprimanded, crying when deprived of material by another child, etc. It does not include crying when the child is hurt.	.28	.28	.50	.00	.40	.13	9.64*
Hyperactivity—Behavior indicating restlessness or giving evidence that the child is under tension. It may be manifested by unusually short span of attention, flitting from one activity to another rapidly, frequently speaking more loudly than necessary, etc. Does not apply to shifting of activity when there are no other signs of tension.	4.00	2.42	3.25	.75	3.40	1.53	1.76
Apprehensiveness—General state of fearfulness, particularly in new situations. Child approaches new situations in tense or rigid manner. Looks to adult for encouragement or clutches adult when facing new situations. Sometimes evidenced only by the expression of the child and withdrawal from the situation. It does not include situations in which there is an observable cause for fear such as sudden loud noises, etc.	9.42	.57	8.75	.50	9.06	.53	5.57*

type, however, we must take account of differences in the variance of the two groups. We are dealing with subjects originally grouped by ratings and there is no *a priori* method of controlling the variances of the two distributions. An application of the F-test indicated that in three of the six patterns the differences in variances are significant below the two per cent level while in three they are not. The three patterns having significant differences in variance are starred in the table. In two of the three patterns showing significant differences in variance, crying and apprehensiveness, there are a large number of low scores in at least one of the groups and this may account for the variance difference. With these reservations we may say that there appears a fairly good indication that we have isolated a number of behavior patterns that differentiate significant groups of secure and insecure children.

Since the individual items tended to differentiate the two groups, we would expect the total of all patterns to yield a highly significant score. The value of "t" for the differences in means of total scores of the two groups is 10.76 which is far beyond the one per cent level of confidence.

The six behavior patterns (crying may be included for the time being until a larger sample has been studied) furnish a method of differentiating secure and insecure children by observational procedures. Observational procedures, however, are relatively expensive and for many purposes, ratings, if fairly reliable and valid, would be helpful. Can the six behavior patterns be used in the form of a rating scale with reliable results? The patterns were cast in the form of a rating scale as follows:

____Voluntary withdrawal from group

This is indicated by an evident preference for playing alone. A child who makes very few, if any, attempts to join a group. This does not apply to a child who withdraws from a group to follow some other interest or in response to observable stimuli.

1. Practically never plays with group
2. Very seldom plays with group
3. Plays with group as much as alone
4. Plays with group more often than alone
5. Plays with group practically all the time

____Non-acceptance by group

This applies to a child who wants to join the group and tries unsuccessfully to do so. Rejection may be physical or verbal. It does not apply to a child who watches a group at play but makes no overt attempt to join it.

1. Practically never accepted
2. Accepted very seldom
3. Accepted as often as not
4. Accepted most of the time
5. Accepted practically all the time

____Bidding for attention of adult

This applies a child who is constantly aware of an adult's presence, and who continually appeals to the adult for attention. This may be manifested in such ways as following adult, seeking unnecessary help, appealing for praise and approval, or any other conduct of a similar nature. It does not apply to a child who asks needed help, or who watches the teacher when she is talking to the group or showing them something.

1. Makes bids for attention very often
2. Makes bids for attention frequently
3. Makes bids for attention as often as not
4. Makes bids for attention infrequently
5. Practically never makes bids for attention

____Crying

Frequent crying in situations which do not ordinarily result in such behavior, such as crying at failure to achieve desired results, crying when reprimanded, crying when deprived of material by another child, etc. It does not include crying when the child is hurt.

1. Cries very often
2. Cries frequently
3. Cries as often as not
4. Cries seldom
5. Practically never cries

____Hyperactivity

Behavior indicating restlessness or giving evidence that the child is under tension. It may be manifested by unusually short span of attention, flitting from one activity to another rapidly, frequently speaking more loudly than necessary, etc. Does not apply to shifting of activity when there are no other signs of tension.

1. Hyperactive very often
2. Hyperactive frequently
3. Hyperactive as often as not
4. Hyperactive seldom
5. Practically never hyperactive

____*Apprehensiveness*

General state of fearfulness, particularly in new situations. Child approaches new situations in tense or rigid manner. Looks to adult for encouragement or clutches adult when facing new situations. Sometimes evidenced only by the expression of the child and withdrawal from the situation. It does not include situations in which there is an observable cause for fear such as sudden loud noises, danger of falling or being hurt by another child, etc.

1. Apprehensive very often
2. Apprehensive frequently
3. Apprehensive as often as not
4. Not apprehensive often
5. Practically never apprehensive

Seventeen children, seven from the two-year group and ten from the three-year group were rated by their respective head teachers and the investigator, all of whom had known the children for some time. A correlation of .88 was obtained. A detailed analysis shows very satisfactory "item by item" agreement as well as agreement in total score. When relatively only approximate data are required, ratings may serve the purpose.

A group of five behavior patterns have been identified which differentiate significantly (within the limits of available statistical technics) a group of "secure" and "insecure" preschool children. A sixth pattern was tentatively retained pending larger sampling. A characteristic of the patterns is the inclusion of the conditions under which the behavior takes place. From theoretical analyses this seems important in developing methods of measuring hierarchy of motives. The behavior patterns may be used in observational procedures or in the form of a rating scale for measuring relative security or insecurity in preschool children.

STUDIES OF ELEMENTARY SCHOOL CLASSES IN ACTION

II. Pupil Participation and Aspects of Pupil-Teacher Relationships

ARTHUR T. JERSILD, BERNARD GOLDMAN
and CATHERINE L. JERSILD
Teachers College, Columbia University

and

JOHN J. LOFTUS
Board of Education, New York City

This paper reports results obtained in one of two special studies that were conducted in connection with a series of investigations designed to evaluate the "activity program" in the New York City Elementary Public Schools.* This educational experiment has been described in several earlier publications (see bibliography). The experiment involved the introduction of newer, "progressive" educational procedures into a sampling of about seventy New York City elementary schools. As a part of the evaluation project, numerous studies have been made, or are still in progress, in which various research procedures have similarly been applied to "activity" schools and to paired "control" schools that have adhered to more traditional practices.

The study here reported was conducted in a limited sampling of classes belonging to four pairs of schools, located in four of the five boroughs of New York. Each pair included an "activity" and a "control" school.*

The first of the two investigations was designed as a comparative study of what actually was happening in the various classes from the time when the children arrived in the morning until they left in the afternoon. This study, entitled, "I. The Working Day in Activity and Non-Activity Schools," was published in the June, 1941, number of this Journal.

In the present study the method of direct observation was used to investigate a number of aspects of the classroom situation, including the amount and nature of participation by individual pupils. This study was an outgrowth, in part, of an investigation initiated by the writer two years ago.** In this preliminary study, it was noted, among other things, that classes under different teachers, and with varying enrollments, had one outstanding characteristic in common, namely the fact that a few children in each class tended to monopolize a large proportion of the discussion.

It also appeared in this preliminary investigation that the tendency on the part of a few children to monopolize recitations and discussions became more pronounced when children themselves were directing the work of the class than when the teacher was directly in charge. These findings, as well as others that will be alluded to later, raised questions that seemed to be quite pertinent to the evaluation project of which the present study is a part, especially in view of the fact that the activity program is designed, at least in part, to promote democratic procedures in the classroom.

Other features of the study afforded additional bases for comparisons and contrasts between the activity and the control classes.

PROCEDURE

The method of direct observation was employed.

The aim of the observers was to note responses by individual pupils and by the teacher which tied in with the class project, and which could be seen or heard by pupils who were giving attention to the work at hand, as distinguished from private interchanges between members of the class, or private undertakings by individual pupils.

* The evaluation was conducted with the aid of workers of the Work Projects Administration, Research and Records Program District Office, Research Survey and Federal Data Section, Official Project No. 65-1-97-21 WP 10 Sub 1.

* The authors gratefully acknowledge the helpfulness shown by the principals and teachers in the schools included in this study, and the conscientious services rendered by the workers who observed in the classrooms.

** Jersild, A. T., Meigs, M. F., and Brown, L. S.: *A Study of Elementary Classes in Action*, Unpublished (New York: Teachers College, *Columbia University, 1939*).

Thus, in a class period devoted to history, any answer by a pupil to a question raised by the teacher, or a report to the class, or a query, comment, or voluntary item of information contributed by a pupil would be entered in the record. A whispered comment, inaudible to the group, or any private or independent act or verbal expression by an individual pupil was not recorded. However, if such behavior by a pupil evoked a comment by the teacher in the hearing of the class as a whole, or if the attention of the class was called to the behavior by the teacher or by another pupil, the comment or notice was included in the record. Further, if an individual comment, whatever the intention behind it might be, was audible to the group and, in effect, functioned as a contribution to the discussion, or patently influenced the work of the class in one way or another, it was noted in the record.

The scope and content of the observational records can be judged from the outline below of the categories that were used in analyzing the data. This list of categories was devised on the basis of an extended period of preliminary observation, including simultaneous but independent observations by two workers whose records were subsequently compared, item by item.

It will be noted from the list below that an effort was made to draw a distinction between certain types of contributions by the pupils, certain features of interchanges between the teacher and the pupils, and between members of the class. Pupil contributions of a recitational nature are distinguished from contributions of a more self-assertive or original character. In the interplay between the teacher and pupils, a distinction is made between (1) attentions of a more or less "neutral" sort, such as being questioned, called upon to recite or directed to perform a certain task; and (2) "positive" attentions (such as praise, being singled out for special commendation, etc.); and (3) "negative" attentions (such as reprimands or rejection of an answer given by a pupil).

CATEGORIES USED IN ANALYZING PUPIL-PARTICIPATION RECORDS

Pupil Contributions of a Recitational Nature in Response to Teacher Questioning, Direction or Prompting

A1. Brief answer—spelling a word, giving seriatim answers in arithmetic drill; also, "yes," "no" on any question; also, single word or phrase or statement of fact in response to direct questioning; giving definitions of common words; etc.

A2. Partial or incomplete or tentatively neither accepted or rejected answer in a series of promptings. (Give one tally to each prompted reply, however brief.) Also, rejected answer, followed by prompting. If answer, following by prompting, is not specifically rejected, do not tally as rejected in reject column, since prompting always implies rejection. However, if there is a specific reprimand or rebuke, tally this elsewhere. Also include unsatisfactory answer; also answer to wrong question—giving answer to problem 5 instead of problem 4, showing confusion or inattention; also answers in terms of a different question or proposition.

A3. Answer of an expository sort, not such as spelling a word, or giving answer to an arithmetic example, but bringing out points or an elaboration in response to direct questioning, also, stating fact plus reason, explanation of fact plus elaboration, of relationship, etc., in response to questioning or direction by the teacher or pupil chairman.

A4. Holds floor for a relatively extended statement of report embodying several contributions of the sort that might receive separate tallies under (3).

A5. Reads aloud, rehearses play.

A6. Demonstrates, illustrates rather than recites; points to map, writes or draws on board, etc.

Pupil Contributions of a Lagging or Retarding Nature

B1. Reports inability to respond when called on—does not have answer to example; has no answer to question of fact or opinion; says, "I don't know."

B1x. Retarded answer, definitely slowing class progress. Teacher waits.

B2. Asks question on point majority are informed about, lags behind, retards progress. (Includes positive statement to same effect, e. g. "I did it entirely differently.")

B3. Asks incidental questions regarding procedure (what kind of paper to use, etc. Says, "Wait a minute!" to teacher during dictation.)

B5. Recites or offers contribution after subject is closed—tries to reopen closed matter.

B6. Irrelevant contribution, beside the point.

Pupil Contributions of a Self-Initiated, Self-Assertive, Critical, Original, Questioning, or Supplementary Nature (Not in response to direct questioning of a recitational nature)

C1. Volunteers supplementary information, suggestions, or questions.

C1x. Makes running comment—remarks made by a child when teacher is speaking or another child is reciting. (Not actual

contributions that advance the main topic, nor yet of an irrelevant nor egocentric nature.)

C2a. Defends own ground against criticism or attack—*work.*

C2b. Defends own ground against criticism or attack—*conduct.*

C3. Praises another's work or contribution. (Do not count here if the comment is in response to a definite teacher request for exposition of good points in another pupil's contribution.)

C7. Egocentric remark; offers unsolicited opinion; expresses attitude; offers personal experience, not generalizing the point.

C8. Thrusts self forward, tries to ask question, to horn in; begins or tries to recite without solicitation or apparent welcome.

C9. Is insistent on making contribution or adding to record an item of personal interest; insists on contributing in spite of contrary suggestions or obstructions.

C3x. Approves teacher—expresses approval, enthusiasm, delight, gratification regarding teacher's practices, including pleasure in privileges granted, assignments, explanations or help offered. (This includes the "OH!" of sudden enlightenment.)

C4. Queries or tries to elicit further information, etc.; constructive or helpful response to pupil report, etc.

C4x. Child directs or instructs another child.

C5. Sympathetic laughter in response to pupil, group or teacher.

C6c. Child excuses another child, explains or condones error in another.

C7. Criticizes, attacks, shows weaknesses, voices objection, counter-opinion.

C8. Assumes role or disciplinarian.

C9. Tries to direct class.

C9x. Child bosses another child—gratuitous direction of a coercive nature, as distinguished from simple directions. (C4x).

C10. Tries to direct teacher or objects to or criticizes teacher; voices opposition, is refractory, groans and shows resistance to teacher requests.

C11. Expresses contempt or disdain, or general negative comment regarding the accomplishments, work or abilities of another pupil.

C12. Derisive laughter at or with pupil, group or teacher.

C14a. Criticizes own work; expounds own error; constructive, self-evaluation; admits own shortcomings; admits improvement.

C14b. Criticizes own conduct.

C15. Child defends child.

Pupil as Recipient of Pupil Reactions: Positive

D3a. Is recipient of pupil approval regarding *work.*

D3b. Is recipient of pupil approval regarding *conduct.*

D5. Is recipient of sympathetic laugh from pupils.

Pupil as Recipient of Pupil Reactions: Negative

D1a. Is criticized or receives expression of individual disapproval for answer, point of view, etc.—*work.*

D1b. Is criticized or receives individual expressions of disapproval for *conduct.*

D2a. Class expression of disapproval regarding an individual's *work*—includes obvious signs of impatience, boredom, sighs, etc., when child lags, takes special time; gives wrong answer.

D2b. Class expression of disapproval regarding an individual's *conduct.*

D6. Is recipient of derisive laugh from pupils.

Attentions from Teacher: Neutral

(All contributions in this group are by the teacher and represent items credited to the child to whom they were directed)

R2. Teacher gives help on solicitation, including pausing and helping child who has asked lagging question.

R4x. Miscellaneous mention (neither favorable nor unfavorable).

R9c. Admonition or criticism of work or conduct stemming from too much zeal or desire to be heard. (Group answers sotto voce when teacher tries to draw answer from laggard.)

R14. Pupil-teacher discussion — content not manifest.

T1. Teacher makes brief interrogatory remark; calls on child for answer, plus all non-interrogatory remarks which aim to draw out facts, information, etc.

T2. Teacher makes prompting remarks—(first in series is always T1 or T5).

T3. Brief expository remarks.

T4. Expository remark of paragraph length or more.

T4.* Expository remarks—several minutes in length.

T5. Brief directions—"Take out books," etc.

T6. Demonstrates, shows how a thing is done.

T6x. Reads, writes, dictates, plays phonograph, leads in a song.

T7. Questions or comments of a nature *enlisting participation,* sharing of work or solution, e. g., "How many agree?" "I think so too"; "Let's show B___ how to do this."

T8. Moralizes, sermonizes, exhorts.

T9. Claps or gives other signal for quiet—"Are you ready now?" "I'm waiting." "Every eye on me."

T10. Positive or negative suggestion or cajolery—"This is easy." "Now we'll find out how smart you are!"

T11. Gives reasons in support of stand or decision—"May I please finish so that those people who are ready can go ahead?"

T12. Interrogatory or expository remarks giving child or children opportunity to express likes or dislikes. "How many like

* All items relating to attentions from the teacher are tallied to the credit of the child or children to whom the attentions are addressed.

the window up?" "Who likes that word?" Chance given to pupils to express what they like.

T13. Interrogatory or expository remarks, giving child or children decision or choice. "What shall we do about the noise?" "What shall I use for a divisor?" (Do not count items in which question is a courteously phrased direction, e. g., "Mary, will you be willing to sit here?") Any chance to do what they like to do.

T15. Offers or gives unsolicited help or attention.

T16. Admits fallibility.

T16x. Admits ignorance, "I don't know, we'll have to look that up," etc.

Attentions from Teacher: Positive

D4ax. Child is approved, applauded by group on teacher solicitation.

R1. Silent acknowledgment, or passive or casual recognition and acceptance of child's contribution. Score in this category when child gives answer to question or volunteers information and teacher says, "Right," or "All right," or picks up the answer in her discourse.

R3. Receives smiling acknowledgment, understanding glances, encouragement, "Thank you!" Tally to child's credit also if he is conspicuously noticed as the only one, or one of few, who raises his hand when asked, who has the right answer, who is finished, etc.

R4. No special word of praise but name singled out, put on board or pupil asked to stand up, as one who has done superior work.

R5a. Brief praise—"good," "fine," etc.—*work.*

R5b. Brief praise—*conduct.*

T14. Teacher acknowledges individual handicap, excuses child—recognizes child may be behind because absent, etc.

T20. Teacher apologizes to child.

Attention from Teacher: Negative

D1ax. Child is criticized or receives individual expression of disapproval from teacher for *conduct.*

D2ax. Is criticized, etc., on teacher solicitation (by group).

R2x. Is denied help on solicitation. ("You do what you think best.")

R4c. Unfavorable recognition or mention (not a verbal reprimand). Slowing action of child recognized as teacher pointedly watches and waits.

R6. Indefinite reprimand. Teacher questions child, calls his name, in a manner suggesting he was lagging or misbehaving. (If child's conduct was observed, record elsewhere.)

R7. Answer passed over, not accepted or rejected in so many words, but substantially rejected.

R10. Rejected ("wrong") corrected ("It should be 144, not 142"), or cautioned against error ("Be sure of your facts," etc.).

R11. Rejected or corrected or cautioned plus special attention of negative sort. Stand up because of error; name put on board as one in need of further study or help.

R12a. Is pointedly ignored or restrained—*work*, contribution, answer, etc. Includes seeking individual attention and told it cannot be given; also asked to wait until another finishes.

R12b. Is pointedly ignored—conduct.

R13a. Rebuke or reprimand—*work*, e. g., "I'm surprised at you!" "That was careless, wasn't it?" (Includes turn of phrase or tone denoting impatience or annoyance with child.)

R13b. Rebuke or reprimand—*conduct.* "Be quiet!" etc., (Relating to specific behavior in progress, as against reprimands or admonitions applying to future conduct or conduct in a general sense.)

R15. Is punished.

T18x. Remark or act of teacher provokes derisive laugh directed against a child.

R8. Admonition regarding *aspect of work*, calculated to improve upon some weakness in child's manner of presentation, but not reflecting on quality of answer and not precisely in the nature of a reproof for behavior—"Speak louder," "More distinctly," "Can't hear you," "Write larger," "Stand so we can see part of the map," "Stand up straight," etc.

R9a. Admonition regarding *work*—"Be careful, think before you speak."

R9b. Admonition regarding *conduct*—"Will not call on wrigglers," "Let's be quiet while we do this."

Pupil Response to Teacher

T17. Teacher receives sympathetic laughter, children laugh at remark or act.

T18. Teacher receives derisive laughter. (Score here only if laugh is directed against teacher; not if teacher says something that makes children laugh derisively at another child.)

Population Studied

Eight activity and eight control classes, representing four pairs of schools, were included in the study. The classes represent the third, fourth, fifth and sixth grades.

Extent of Observations

Each class was observed on three or four different days, but the total amount of time devoted to active observation varied somewhat from class to class, depending upon the nature of the program and the occurrence of field trips and other interruptions. Moreover, the observers took time off at intervals while they were in the classroom, since the work of watching, listening and recording is rather

exhausting. The amount of time spent in active observation in the various classrooms ranged from 250 to 478 minutes with a median of 397 minutes, or over six and a half hours.

Measurements of Reliability

The reliability of the observers when utilizing the procedures employed in this investigation was measured in the earlier study, referred to above, by the method of item by item comparison of records obtained by workers who simultaneously but independently recorded the same behavior. The percentage of agreement in various comparisons between four different observers ranged from 76 to over 90. These percentages were computed on the basis of sub-categories more numerous and refined than those represented in the tables that follow. The agreement in terms of the larger groupings of behavior, as represented in Table II and following, was over 90%.

In the present study, similarly detailed analyses of observer-reliability were not undertaken except during the preliminary period. Various checks and safeguards were, however, applied.* First, all of the observers devoted several days to a period of training, under the direction of one of the four observers who participated in the original study in which this method was employed and whose reliability as an observer had been established. Records that were obtained by the present workers while observing simultaneously with this previously-trained worker were compared, item by item, and discussed in detail, before the regular schedule of observations was begun.

Further, as a subsequent rough measure of reliability, each of the four observers was paired with one of the others during half-hour periods of simultaneous observation. Each observer's record was then separately analyzed, tallies were entered for each pupil, and thereafter the correlation between the pupils' participation scores, as represented by these independent analyses, were computed. The coefficients thus obtained ranged from .37 to .83 in the case of the activity classes and from .24 to .79 in the case of the control classes. The respective median coefficients were .62 and .66. These coefficients certainly are not high, but they are considerably higher than the author would have predicted, for several reasons: first, each coefficient represents only a half-hour sampling of behavior; second, they are based upon comparisons between pairs of workers, one of whom, and sometimes both of whom, were completely unfamiliar with the group that was being observed. It seems reasonable to conjecture that if the sampling had been larger, and if observers whose records were being compared had been fully familiar with the groups, the coefficients would have been substantially larger.*

Reliability of the Sampling

To obtain an indication of the reliability of the sampling, scores were separately computed on the basis of the first two days of observation and the second two days of observation. Since the results as far as individual pupils are concerned are of interest mainly in connection with the categories that deal with the contributions made by the pupils, these reliabilities were computed only for the two major categories of pupil participation (recitational contributions and self-initiated or self-assertive contributions).

The coefficients obtained from the correlation of the results of two days against the remaining two days are shown in Table I. It will be noted that the coefficients relating to Category C (self-initiated and self-assertive contributions) were highest in the case of the activity schools, ranging from .57 to .89 with a median of .80. The corresponding coefficients based on the results for control classes ranged from .15 to .77 with a

* As can be recognized, it requires quite an extended period of experience with a class as large as those here involved before an observer is thoroughly familiar with the names, appearance, voices and mannerisms of the various members. Until such familiarity is acquired, errors in observation are influenced to a large extent by factors which time and experience will remedy. By the time the present study was begun, each observer was already familiar with his classes, by virtue of having spent four entire days in each class in connection with "Working Day" observations. In order to function adequately as a check observer, another worker would have to spend a good deal of time in classes not familiar to him. Such paired observations were dispensed with, since there was a pressure for time. However, various expedients for measuring and safeguarding the accuracy of the observations were resorted to, as described above.

* From the point of view of methodology, an incidental item may be noted here, namely, that the correlations tended to be *lower* when based upon the results of one observer who was thoroughly familiar with the group and of another who was quite unfamiliar than when based upon results obtained by observers both of whom were unacquainted with the group. A high correlation in the latter case perhaps reflects the fact that both observers make similar errors or omissions quite as much as it can be interpreted to mean that the method of observation is sound. This state of affairs adds further point to comments made in another connection concerning shortcomings of the use of the correlation method in determining observer-reliability. (Jersild, A. T. and Meigs, M. F.: "Direct Observation as a Research Method." *Review of Educational Research*, IX, (1939) 472–482; 597–599.)

TABLE I

RELIABILITY OF THE SAMPLING, IN SELECTED CLASSES, OF BEHAVIOR OF INDIVIDUAL PUPILS IN CATEGORIES A (RECITATIONAL CONTRIBUTIONS), C (SELF-ASSERTIVE AND SELF-INITIATED CONTRIBUTIONS), AND CATEGORIES A AND C COMBINED IN THE VARIOUS CLASSES THAT WERE OBSERVED DURING FOUR DAYS. FIGURES ARE BASED UPON CORRELATIONS BETWEEN TWO SETS OF TWO DAYS OF OBSERVATION

(The Values Given for Number of Children in Each Class Represent Only the Pupils Who Were Present Throughout the Four-Day Period of Observation)

School & Grade	No. of Children	Category A	P.E.	Corrected r	Category C	P. E.	Corrected r	Category A & C	P. E.	Corrected r
Activity										
P.S. I–4B[3]	36	.50	.08	.67	.40	.09	.57	.66	.06	.80
P.S. II–4B[3]	24	.20	.13	.33	.80	.05	.89	.68	.07	.81
P.S. III–3A[4]	27	.05	.13	.10	.67	.07	.80	.58	.09	.73
P.S. IV–5A[1]	32	.15	.12	.26	.61	.07	.76	.57	.08	.73
P.S. V–3B[3]	23	.31	.13	.47	.72	.07	.84	.61	.09	.76
Control										
P.S. I–6B[3]	22	—.38	.12	—.55	.62	.09	.77	—.003	.14	—.01
P.S. II–3A[4]	21	.24	.14	.39	.08	.15	.15	.05	.15	.10
P.S. III–5A[5]	26	.46	.10	.63	.56	.09	.72	.63	.08	.77
P.S. IV–6B[3]	23	.39	.12	.56	.35	.12	.52	.41	.12	.58
P.S. V–3B[4]	18	.25	.15	.40	.19	.15	.32	.24	.15	.39
P.S. VI–5B[3]	27	.61	.08	.76	.63	.08	.77	.71	.06	.83

median of .62. The coefficients representing Categories A and C combined ranged from .73 to .83 in the activity classes and from —.01 to .83 in the control classes.

AGREEMENT ON CLASSIFICATIONS

The categories used in the analysis of the pupil-participation records were checked by the procedure of comparing, item by item, the results obtained when workers independently analyzed portions of the data. Several such comparisons were made in the present study, as well as in the earlier study referred to above. The agreement between independent workers with respect to the major categories represented in the tables that follow ranged from 70 per cent to over 95 per cent with a median agreement of over 90 per cent.

RESULTS

Table II shows the results for the eight activity and eight control classes in terms of total frequencies of various types of behavior under various general headings.

RELATIVE FREQUENCY OF RECITATIONAL AND SELF-INITIATED CONTRIBUTIONS

It appears from Table II that the frequency of self-assertive or self-initiated pupil contributions was considerably higher both in relative and in absolute terms in activity classes than in control classes. In the activity classes the ratio of "recitational" items to self-initiated or self-assertive items was 1.52 as contrasted with a corresponding ratio of 3.63 in the case of control classes. When totals for these two categories are computed, it appears that 40 per cent of the total in the activity classes was of the self-initiated variety, while the corresponding value in the control classes is only about 22 per cent.

It can be noted from Table II that this difference appears not only when results are computed for all classes combined, but also appears quite consistently, to varying degrees, when comparisons are made between matched classes.

Perhaps the difference here noted is no more than one would expect, yet it shows quite impressively that the procedures differ. In the activity classes there is considerably less catechizing of a piece-meal, question-and-answer type than is the case in the control classes, and there is considerably more give and take of a discussional nature.

POSITIVE AND NEGATIVE PUPIL-PUPIL REACTIONS

From Table II it can also be noted that the children in activity classes have more interchanges by way of criticism and praise of one another's work. The "negative" pupil-reactions in the activity classes outnumber

TABLE II

TOTAL TALLIES, FOR EACH CLASS, OF VARIOUS FORMS OF PUPIL PARTICIPATION AND OF ATTENTIONS FROM TEACHERS, INCLUDING, ALSO, FOR EACH CLASS, (1) THE RATIO OF RECITATIONAL TO SELF-ASSERTIVE OR SELF-INITIATED PUPIL CONTRIBUTIONS, AND (2) POSITIVE TO NEGATIVE TEACHER RESPONSES TO PUPILS

A = activity, C = control Schools & Grades		No. of min. observed	Recitational Contributions	Self-Initiated & Self-Assertive Contributions	Ratio A to C	"Lag" Items	Negative Pupil-Pupil Reactions	Positive Pupil-Pupil Reactions	Neutral Teacher Attentions	Positive Teacher Attentions	Negative Teacher Attentions	Ratio of Positive to Negative	Pupil Approval of Teacher	Pupil Disapproval of Teacher	No. of Miscellaneous Duties
A	4B[1]	425	285	193	1.48	8	55	13	311	12	61	.20			38
C	4B[2]	335*	215	53	4.06	11	0	9	493	42	107	.39	2		68
A	6A[3]	330*	180	58	3.10	17	3	0	578	32	122	.26			66
C	6B[3]	478	448	44	10.18	23	10	0	920	51	82	.62			126
A	4B[2]	394	285	146	1.95	28	40	13	456	11	87	.13			35
C	4B[2]	270*	134	49	2.73	15	8	9	197	10	59	.17			32
A	6A[3]	250*	176	159	1.11	6	49	8	175	9	36	.25			26
C	6B[3]	350	179	93	1.92	21	11	12	340	14	50	.28			16
A	3B[2]	404	468	331	1.41	29	51	15	727	68	151	.45			21
C	3B[4]	464	333	44	7.57	31	15	0	774	5	106	.05	4		58
A	5A	327*	248	158	1.57	47	27	3	562	38	163	.23			43
C	5B[2]	432	831	296	2.81	57	17	3	1161	110	125	.88			42
A	3A[4]	400	413	227	1.82	83	36	2	779	28	395	.07			72
C	3A[4]	400	625	121	5.17	104	14	3	1032	90	348	.26			154
A	5A[1]	380	281	266	1.06	5	45	7	574	5	28	.18			27
C	5A[5]	430	472	192	2.46	43	30	8	632	19	52	.37	1		93
Total: all activity classes			2336	1538	1.52	223	306	61	4162	208	1043	.195			328
Total: all control classes			3237	892	3.63	305	105	44	5549	341	929	.37			589

*3 days of observation only.

corresponding responses in control classes to the extent of almost three to one. "Positive" reactions—representing praise of one pupil by another, laudatory comments, etc.—also are relatively more frequent in the activity classes, but the total number of such responses is relatively small compared with the number of negative responses, and the positive items in the activity classes outnumber similar items in the control classes in a ratio of 1½ to 1.

Praise and Reproof from the Teacher

It is interesting to note that in no class, either in the control or the activity groups, do positive attentions from the teacher outnumber negative items as defined in the set of categories described earlier in this paper. In all classes, the children, as a group, are reprimanded or otherwise reminded of their shortcomings more frequently than they are praised or commended.

In the activity classes, the cumulative total of "negative" attentions from the teacher is 1043, as contrasted with 203 "positive" items. In other words, occasions on which children are reproved, pointedly ignored, reminded of shortcomings, admonished against the repetition of past errors, or specifically reminded of error, outnumber expressions of praise in the ratio of more than five to one. In the control classes, the corresponding cumulative totals are 929 and 341, representing a negative-positive ratio of somewhat less than three to one.

In this connection, it may be of passing interest to compare the results in the present study with the findings in the earlier study of four classes mentioned above. In this earlier study two of the four classes showed a higher incidence of positive attentions from the teacher. In the third class the frequencies were about the same, but in the fourth class, which corresponded more closely in size to the classes in the present study, the negative items outnumbered the positive items. The higher incidence of negative items appeared to be influenced, to a large extent, by the size of the class. In the large class in question, the teacher was very genial and seemed to be very popular with the pupils, even though her negative comments quite exceeded her positive comments.

The fact that the ratio of negative to positive attentions from the teachers is higher in the activity than in the control classes is interesting, but it is a finding that perhaps might well have been expected. To the extent that greater freedom and informality prevail, there would be more occasions to call individual pupils to account than would be the case when all children are regimented as a group.

The fact that there is a preponderance of negative over positive responses from the teachers in both groups of schools is of passing interest in view of findings, by Thorndike and others, that indicate that rewards tend to be more effective than punishments as incentives to learning. It is possible that efforts by the teachers to use a higher proportion of positive remarks would have a salutary effect. It may be pointed out, however, that the "rewards" received by the pupils are not limited to words of praise or commendation such as those here classified as "positive" responses. In a great many instances, the pupils undoubtedly receive the equivalent of a reward through the experience of success that they have when they answer correctly or contribute effectively to group work or class discussions, even though they win no tangible rewards or expressions of praise. In some instances, no doubt, the very fact that a teacher accepts a pupil's contribution without comment or criticism, may, in effect, represent a form of commendation. If effects of this sort, together with subtle forms of commendation that a teacher might communicate by means of facial expressions or other signs, could have been taken into account it is quite possible that the ratio of "positive" to "negative" items, as here reported, would be considerably changed.

Distribution of Participation by Individual Pupils

The findings, as shown in Table II, indicating that pupils in activity classes have more freedom to make contributions of a self-initiated or self-assertive nature, to volunteer questions, comments, criticisms and supplementary information during the course of class discussions, raise a question which, in the writer's judgment, is quite important from the point of view of the evaluation of the activity program. Under this regime of greater freedom, to what extent are the op-

portunities for participation utilized by all members of the class?

The tables which follow deal with this question. These tables show comparisons between activity and control classes on the following points: the range in number of contributions made by the individual pupils in each of the classes during the days when the classes were observed; the median number of contributions; the per cent of the total contributions that were made by the pupil who contributed most often and by the three pupils who contributed most often.

The tables also show the number of children from the *low end* of the distribution of scores whose combined contributions equaled the contributions made by the single child who talked most. (A value, say, of fifteen in this computation means that the contributions made by the child who talked most were equal in number to the cumulative contributions made by the fifteen least talkative members of the class.) To find this value a table was constructed which showed the total number of times each child contributed; the subjects were arranged in order from the child who contributed most to the child who contributed least; then, beginning at the bottom of the table, a tally was made upward, through the distribution, until the cumulative tally equaled the score of the child who stood at the top of the list.

In treating the results from this point of view, use was made only of those children who were present during all of the days when the class was observed. This means that the class population included in these computations was smaller than the total enrollment of the class.

Table III shows the results for each activity and control class; Table IV shows the medians of the values for the 8 classes in each of the two groups. (In computing these medians, it will be noted, the classes were treated as subjects; the medians were not computed on the basis of a distribution representing each of the individual children in various classes.)

The entries showing the median number of contributions per child are in keeping with the results already noted above. They indicate that in the activity classes there were considerably more contributions of the self-initiated, self-assertive variety than in the control classes and that children in the activity classes made fewer contributions of a recitational or brief question-and-answer variety. In spite of these differences, however, it appears that there is a rather high degree of similarity between the two groups with respect to the manner in which these contributions were distributed between members of the class. In the case of recitational items the child who spoke most often contributed approximately 9 per cent of all items that were recorded in both the activity and control groups. Similarly the median of the values for the separate classes indicates that about 1 per cent of the recitational items were contributed by the children who talked least.

The results shown in Table III and IV are quite arresting. They indicate the degree to which inequalities of participation prevail. They suggest, among other things, that participation in group discussion shows a decidedly skewed distribution rather than the approximately "bell-shaped" distribution that one frequently finds in measurements of ability, physical dimensions, and the like. They also suggest that the label "free discussion" is largely a misnomer, for when there is opportunity for "free" discussion (in the sense that there is a chance to offer contributions such as those classified as self-initiated, self-assertive, voluntary, critical, the like, in the present study) a few members of the group proceed to monopolize a large part of the discussion, with the result that many members of the group seldom or never are heard from.

Even in a class under more formal direction it is true—as can be seen from the results in the first section of each of the two tables—that some children contribute considerably more than do others, but the disproportion is not nearly so large. The child who recites most often in response to teacher solicitation makes about 9 per cent of the total recitations, but this value is still considerably smaller than the corresponding value of about 16½ per cent which is found in connection with responses of the self-initiated and self-assertive variety.

The third section in each of the tables offers some information concerning the distribution of contributions in the recitational and the self-assertive categories combined. In these sections it can be seen that, in some comparisons, the differences between the control and the activity classes is somewhat

TABLE III

DISTRIBUTION OF PARTICIPATION BY PUPILS IN ACTIVITY AND CONTROL CLASSES. RESULTS ARE SHOWN FOR (A) RECITATIONAL CONTRIBUTIONS AND (C) SELF-INITIATED, SELF-ASSERTIVE, VOLUNTARY AND SUPPLEMENTARY CONTRIBUTIONS

	I. Activity Classes							
School	A	B	C	D	E	F	G	H
Grade	4B[3]	6A[3]	4B[3]	6A[3]	3B[3]	5A	3A[4]	5A[1]
No. of pupils present throughout the observation period	24	28	36	44	23	24	27	32
A. Recitational contributions								
Range of scores	1–25	1–9	0–21	0–12	5–41	2–18	7–21	0–17
Per cent of total contributed by pupil with highest score	10%	6%	10%	7%	10%	10%	6%	8%
Per cent of total contributed by pupil with lowest score	.4%	1%	0	0	1%	1%	2%	0
No. of children at lower end of distribution whose cumulative scores were equalled or exceeded by the child who contributed most	7	4	14	14	5	6	3	9
Total contributions made by the three children with highest scores	69	27	55	31	93	49	59	45
Per cent of total contributed by the three pupils who contributed most	28%	17%	27%	19%	24%	27%	17%	21%
Per cent of contributions made by pupils in highest quartile	47%	37%	57%	50%	40%	46%	36%	45%
Median number of contributions per child	9.0	5.5	3.5	3.3	17.8	7.1	14.2	6.6
C. Self-initiated, supplementary, self-assertive contributions								
Range of scores	0–33	0–10	0–12	0–15	1–44	0–18	1–45	0–35
Per cent of total contributed by pupil with highest score	20%	21%	10%	10%	15%	15%	22%	18%
Per cent of total contributed by pupil with lowest score	0%	0	0	0	.3%	0%	4%	0
No. of children at lower end of distribution whose cumulative scores were equalled or exceeded by the child who contributed most	15	19	16	20	10	14	16	20
Median no. of contributions per child	5.1	1.0	2.4	2.4	11.1	3.5	2.0	3.5
Total contributions made by the three children with highest scores	78	20	33	37	103	50	99	68
Per cent of total contributed by the three pupils who contributed most	46%	42%	28%	26%	36%	41%	49%	35%
Per cent of contributions made by pupils in highest quartile	63%	71%	62%	61%	57%	63%	69%	67%
A. and C. combined								
Range of scores	1–51	1–19	0–28	0–20	8–65	2–34	10–62	0–48
Per cent of total contributed by pupil with highest score	12%	9%	9%	6%	10%	11%	11%	12%
Per cent of total contributed by pupil with lowest score	.2%	0	0	0	1%	.66%	2%	0
No. of children at lower end of distribution whose cumulative scores were equaled or exceeded by the child who contributed most	9	6	12	12	5	8	5	11
Total contributions made by three children with highest scores	130	47	81	55	179	91	149	102
Per cent of total contributed by three pupils who contributed most	32%	23%	25%	18%	26%	30%	27%	25%
Per cent of contributions made by pupils in highest quartile	50%	44%	55%	52%	45%	52%	46%	50%
Median no. contributions per child	14.5	6.5	5.5	5.5	27.1	9.4	15.8	10.3

TABLE III—Continued

	II. Control Classes							
School	Ac	Bc	Cc	Dc	Ec	Fc	Gc	Hc
Grade	4B[3]	6B[3]	4B[3]	6B[3]	3B[4]	5B[3]	3A[4]	5A[5]
No. of pupils present throughout observation period	30	23	27	22	18	27	21	26
A. Recitational contributions								
Range of scores	0–16	7–26	0–8	2–14	2–31	4–61	11–34	4–26
Per cent of total contributed by pupil with highest score	10%	8%	7%	9%	12%	10%	8%	9%
Per cent of total contributed by pupil with lowest score	0	2%	0%	1%	1%	.7%	3%	1%
No. of children at lower end of distribution whose cumulative scores were equaled or exceeded by the child who contributed most	10	4	6	4	6	9	3	5
Total contributions made by the three children with highest scores	40	71	22	36	81	177	91	63
Per cent of total contributed by the three pupils who contributed most	24%	21%	19%	22%	32%	29%	21%	22%
Per cent of contributions made by pupils in highest quartile	52%	38%	41%	39%	39%	55%	33%	37%
Median no. of contributions per child	4.6	14.3	4.5	7.5	14.7	18.9	19.8	10.5
C. Self-initiated, supplementary, self-assertive contributions								
Range of scores	0–8	0–6	0–12	0–10	0–6	0–36	0–9	0–18
Per cent of total contributed by pupil with highest score	21%	18%	27%	14%	18%	15%	11%	15%
Per cent of total contributed by pupil with lowest score	0	0%	0%	0%	0	0	0%	0%
No. of children at lower end of distribution whose cumulative scores were equaled or exceeded by the child who contributed most	21	13	18	11	13	13	7	11
Total contributions made by three children with highest scores	18	14	18	28	17	87	25	42
Per cent of total contributed by three pupils who contributed most	47%	41%	41%	38%	50%	37%	30%	35%
Per cent of contributions made by pupils in highest quartile	76%	68%	64%	64%	68%	63%	44%	53%
Median no. contributions per child	.7	1.1	1.1	2.3	.5	5.9	3.8	3.8
A and C. combined								
Range of scores	0–24	7–30	0–18	4–24	2–37	6–96	15–40	5–44
Per cent of total contributed by pupil with highest score	12%	8%	11%	10%	13%	11%	8%	11%
Per cent of total contributed by pupil with lowest score	0	2%	0%	2%	1%	.7%	3%	1%
No. of children at lower end of distribution whose cumulative scores were equaled or exceeded by the child who contributed most	12	4	7	5	6	10	3	6
Total contributions made by three children with highest scores	51	83	38	58	90	264	102	98
Per cent of total contributed by three pupils who contributed most	25%	23%	24%	25%	31%	31%	20%	24%
Per cent of contributions made by pupils in highest quartile	53%	40%	43%	43%	39%	56%	31%	40%
Median no. contributions per child	6.0	14.9	5.7	10.4	16.5	28.2	24.7	15.1

TABLE IV

DISTRIBUTION OF PARTICIPATION BY INDIVIDUAL PUPILS: MEDIANS OF RESULTS OBTAINED IN ALL CLASSES

	Activity	Control
No. of classes	8	8
A. Recitational contributions		
Median of median contributions per child	6.85	12.4
Median per cent of contributions made by the most talkative pupil in each class	9%	9%
Median per cent of contributions made by pupils with the lowest score in each class	.7%	1.0%
Median number of children at lower end of distribution in each class whose contributions were equalled or exceeded by the pupil with the highest score	6.5	5.5
Median per cent of total contributions made by the three children with the highest scores in each class	22.5%	22%
Median per cent of contributions made by pupils in the highest quartile in each class	45.5%	39%
C. Self-initiated, supplementary, self-assertive contributions		
Median of median contributions per child	2.95	1.5
Median per cent of contributions made by the most talkative pupil in each class	16.5%	16.5%
Median per cent of contributions made by pupils with the lowest score in each class	0%	0%
Median number of children at lower end of distribution in each class whose scores were equalled or exceeded by the pupil with the highest score	16	13
Median per cent of total contributions made by the three children with the highest scores in each class	38.5%	41%
Median per cent of contributions made by pupils in the highest quartile in each class	63%	64%
A. + C.		
Median of median contributions per child	9.85	15.0
Median per cent of contributions made by the most talkative pupil in each class	10.5%	11%
Median per cent of contributions made by pupils with the lowest score in each class	.1%	1.0%
Median number of children at lower end of distribution in each class whose scores were equalled or exceeded by the pupil with the highest score	8.5	6
Median per cent of total contributions made by the three children with the highest scores in each class	25.5%	24.5%
Median per cent of contributions made by pupils in the highest quartile in each class	50%	41.5%

larger than corresponding differences in the preceding sections of the tables. This greater degree of difference arises in part by virtue of the fact that in the activity classes there is a larger total number of contributions of the self-initiated or self-assertive variety, and as noted above, it is in connection with this type of contribution that the greatest inequalities between individual pupils prevail. It can be noted, for example, that the volume of talk delivered by the median most talkative child in the activity classes exceeds the combined volume of the 8 least talkative children. In the control classes the median most-talkative child out-talks 6 other children combined.

Similarly, it can be noted that, in the activity classes, one-fourth of the children do one-half of the total amount of talking while in the control classes one-fourth of the children do about 41.5 per cent of the talking. The difference here is not large but it does indicate that the disproportions are somewhat larger in the activity classes. It would appear, in other words, that one effect of the greater degree of freedom and informality that prevails in the activity classes is to give freer range to the tongues of talkative children. It is noteworthy, however, that although the disproportion is somewhat larger in the activity classes, the difference between the activity and control classes still is relatively small. In other words, it can not be said that the activity program has primarily had the effect of providing freedom that is primarily preëmpted by only a few children. Neither can it be said, on the other hand, that the freedom provided by the activity

program has introduced relatively more universal participation, at least as far as verbal behavior is concerned.

Quite as striking are similarities between the activity and control groups with respect to the self-initiated and self-assertive types of contributions. In the activity classes the most loquacious pupil contributed from 10–22 per cent of all items; the median of the class values is 16.5 per cent. In the control classes the most loquacious pupils contributed from 11–27 per cent of all items; the median of the class values in this case is also 16.5 per cent. It appears that whether the class situation allows for much or for relatively little pupil participation of this sort, the opportunities that are provided will be preëmpted to a similarly large extent by the most talkative child.

In both the activity and the control groups the median of the values shown for the separate classes likewise show that the least talkative pupil made no contributions whatsoever of this type. Similarly the three most talkative pupils in both the activity and the control groups contributed approximately 40 per cent of all that was said.

One of the entries does show a slight difference: When a tally was made of the number of pupils at the lower end of the distribution whose cumulative contributions were equaled or exceeded by the most loquacious child, the results show that the most talkative child bestrode a somewhat larger number of children in the activity than in the control schools. The difference is not large, however. In the activity schools the range is from 10–20 and the median of the separate values is 16. In the control classes the results for the individual classes show values ranging from 7–21, and the median of the separate class values is 13.

It likewise appears that approximately one-fourth of the children in the various classes did approximately two-thirds of the talking.

It is only fair to state, parenthetically, that the situation that is here found to prevail among children in the classroom is not unique. In ordinary observation of groups of adults we can note that the same disproportion of verbosity prevails. By way of an informal experiment, the writer undertook, over a period of time, to make a record of the behavior of adults in committee meetings, seminars and discussions of various kinds. The results indicated that the tendency of a few individuals to monopolize the discussion was even more pronounced in these adult groups (limited in number) than in groups of children. The results were not extensive enough, however, to provide conclusive comparisons. It is rather interesting to note that a commodity so abundant as hot air should be monopolized so much.

Relation of Participation to Information and Ability

As indicated in the introduction to this study, the results of an earlier investigation indicated that in three out of four classes there was little or no correlation between the amount of talking children did and the amount of information they possessed. In the present study an effort was made to obtain some information concerning the relationship between the competence of the pupils and the extent to which they contributed. Table V shows the correlation between children's scores in the category of contributions that includes self-initiated, self-assertive, voluntary, critical, etc., remarks and the scores of the same children on various tests. The tests listed in the table have been described in earlier publications.* They include tests of ability to apply generalizations in the social studies, of working skills in the social studies, and of knowledge of current events; they also included a battery of tests of reading comprehension and speed, arithmetic computation, arithmetic reasoning, spelling, and language usage. They included also paper and pencil personality tests.

It can be seen from Table V that practically all the correlations between achievement on various tests and amount of self-assertive participation in class discussion are low and unreliable. Most of the coefficients are positive, however.

The median of the coefficients shown by the activity classes in correlations between number of self-assertive contributions and scores in a test of ability to apply generalizations, is .125, which represents practically a zero correlation. The corresponding median coefficient in the control classes is .185, which is slightly higher than the value found in the activity classes, but it still perhaps

* Jersild, A. T., Thorndike, R. L., Goldman, Bernard, and Loftus, J. J.: "An Evaluation of Aspects of the Activity Program in the New York City Public Elementary Schools," *Journal of Experimental Education*, VIII, (December, 1939), pp 166–207.

TABLE V

CORRELATIONS BETWEEN THE FREQUENCY OF PUPIL'S SELF-INITIATED AND SELF-ASSERTIVE CONTRIBUTIONS IN CLASS AND THEIR SCORES ON VARIOUS TESTS

School	Class	Number of Pupils*	Test: Applying Generalizations		Current Affairs		Working Skills		Modern School Achievement**		Personality	
			r	P. E.	r	P. E.	r	P. E.	r	P. E.	r	P. E.
Activity Schools												
I	5A	24	.21	.13	.42	.11	.38	.12	.55	.11	.06	.14
II	5A[1]	32	.04	.12	.21	.11	.31	.11	.29	.12	.03	.12
III	4B[2]	24	.12	.14	.36	.13	.27	.13	.20	.14	.52	.10
IV	6A[3]	26	.11	.13	.53	.10	.10	.13	.12	.13	.22	.13
V	4B[2]	36	.13	.11	.38	.10	.29	.11	.43	.11	.14	.11
VI	6A[3]	44	.26	.10	.18	.10	.04	.10	.21	.12	.18	.10
Non-Activity Schools												
I	5B[3]	27	.16	.13	.28	.13	—.04	.14	—.05	.14	.09	.14
II	5A[5]	26	.21	.13	.38	.12	.35	.12	.44	.12	.05	.13
III	4B[3]	27	—.08	.13	.26	.13	.10	.13	.04	.13	—.02	.12
IV	6B[3]	22	.53	.11	.40	.12	.24	.12	—.03	.17	.05	.13
V	4B[3]	27	.05	.13	—.14	.13	—.12	.13	.11	.14	—.30	.12
VI	6B[3]	23	.52	.11	.56	.10	.23	.14	—.28	.20	—.05	.15

* This represents the number of pupils present on all days when the classes were observed. This number is smaller than the enrollment in each class.

** Median score in a battery of tests including: reading level of comprehension, reading speed and accuracy, arithmetic computation, arithmetic reasoning, diagnostic test in spelling, and language usage.

represents a negligible correlation. It is interesting to note, however, that a coefficient of over .50 is shown in two control classes (both at the sixth grade level), as against a negative correlation of .08 in a class at the other extreme. All in all, the results suggest that perhaps the children who talk most in the control classes are relatively abler than those who talk most in activity classes, but the findings are quite inconclusive on this point.

Knowledge of current affairs shows a higher correlation with talkativeness than do scores in the test of ability to apply generalizations in the activity classes (the median coefficient of correlation in the case of knowledge of current affairs is .37). On the other hand, the corresponding median coefficient in the control classes is only .125. In the activity classes it appears that interests and abilities of the sort that are involved in keeping abreast of current events as reported in newspapers, news broadcasts and magazines are relatively more conspicuous in loquacious pupils than is competence such as that which is measured by the test of ability to apply generalizations. The results do now show to what extent the child who is inclined toward loquacity reads the newspapers as a convenient source of material to talk about, or to what extent greater knowledge of current events is a spur to greater talkativeness.

The correlation between amount of talking and scores on a test of working skills in the social studies is consistently low in all classes in both groups.

The entries representing correlations between talkativeness and scores in the Modern School Achievement Tests are of some interest, since the latter represent a comprehensive battery of standardized tests of competence in various subject-matter skills. The median coefficient in the case of the activity classes is .25; the corresponding value for the control classes is .005. Most of the coefficients are low and unreliable but the direction of the difference is noteworthy. It appears that the more talkative children in the activity classes tend to be relatively more (if only slightly more) competent in subject-matter than the talkative control children, just as the most talkative activity children tend to be relatively better informed on current events than are the talkative control children.

The correlation between talkativeness and scores on the personality test that was administered are low and inconclusive. In the activity classes, all the coefficients are positive, indicating that the more talkative members possessed, if anything, slightly better emotional adjustment than those who talked less. In the control classes three of six coefficients are negative (two of these three are practically of zero value). To the extent that results of a paper-and-pencil test of personality may be regarded as significant, it appears that there is a slightly higher positive relation between personality adjustment and talkativeness in the activity than in the control classes.

All in all, it appears from the findings discussed above that the more talkative members in both activity and control classes tend slightly to surpass the less talkative members in information, academic achievement, and personality ratings. However, the coefficients are, in general, so low that information, academic achievement and personality adjustment cannot be regarded as potent factors in determining the relative amount of talking a pupil will do. Rather, it must be assumed that other factors, including, no doubt, motives associated with unascertained personality traits, are most important.

More detailed information on this matter was obtained in an earlier study. Needless to say, many different motives appeared to operate. In one class, for example, two boys who ranked near the top in amount of talking were rivals, both in the classroom and on the playground. In the same class, the child who talked most was a relatively unpopular youngster whose aim seemed always to be to occupy a prominent place. There was one inveterate talker, in another class, who seemed to be trying, through participation in the class discussion, to compensate for his inability to participate effectively in the out-of-class play activities of his peers. Another child, likewise a very frequent contributor, made his remarks in a rather genial and friendly manner—a large number of his contributions were of an "echo" type, such as repeating, with a complimentary smile, a statement made by another pupil, or adding another point or illustration similar to one just submitted by a classmate, or commenting briefly and favorably about something said by another pupil. Again there were pupils

whose tenacity in re-asserting a statement of fact or opinion appeared to be in the nature of resistant behavior. There also were pupils who were silent but very able and who seemed to be content to let others keep up the flow of words.

Summary

In this study observers in the classroom recorded the contributions made by children during recitations, discussions and other class activities. Records were also made of interactions between the teacher and pupils.

The subjects in the study included the pupils enrolled in 8 activity and 8 control classes representing four pairs of schools. The classes represented the 3rd, 4th, 5th, and 6th grades.

Each class was observed during from three to four days. The observers kept records throughout the day except for intervals when the observer rested and occasions when the class left the room for field trips or for other activities. The amount of time spent in active observation in the various classrooms ranged from 250 to 478 minutes with a median of about 6½ hours.

The data were analyzed in terms of a lengthy list of categories that are described in detail in an earlier part of this report. In the classification a distinction was made between contributions by pupils that were primarily of a recitational or formal question-and-answer variety and contributions of a more self-assertive, self-initiated or supplementary variety, including unsolicited statements of opinion, queries, criticisms, defense of own work or point of view, volunteered supplementary information, etc.

In the analysis of reactions between the teacher and the pupils, a distinction was made between "neutral," "positive" and "negative" remarks or gestures addressed by the teacher to the pupils.

Steps that were taken to measure the reliability of the technique that was used and of the data that were obtained are described in the main body of the report.

The frequency of self-assertive or self-initiated pupil contributions was considerably higher both in relative and absolute terms in activity classes than in control classes. In the former group the ratio of "recitational" items to self-assertive or self-initiated items was 1.52 as contrasted with a corresponding ratio of 3.63 in the case of the non-activity classes. This difference appeared not only when results were computed for all classes combined but also to varying degrees when comparisons were made between matched classes. The results indicate that in the activity classes there was considerably less catechizing of a piece-meal question-and-answer type.

In the activity classes there also were more interchanges between pupils by way of criticism and praise of one another's work.

Negative criticisms or other negative reactions directed toward fellow pupils were considerably more frequent than positive reactions directed toward fellow pupils were con-another, laudatory comments and the like.

In both the activity and the control groups positive items from the teacher (explicit praise, commendation, and the like) occurred considerably less often than did negative items from the teacher (explicit reproof, reprimands, and the like). In the activity classes the total number of "negative" items was 1043 as contrasted with 203 positive items. In the control group the corresponding, respective, totals were 929 and 341. It appears that when a greater degree of freedom prevails there are more numerous occasions on which the teacher feels called upon to correct or admonish individual pupils than is the case when the entire group spends relatively more time on a single project and can therefore be regimented as a group.

There was considerably less equality of participation in connection with contributions of a self-initiated or self-assertive variety than in connection with recitational activities.

In both activity and control classes the median values indicate that 9 percent of all *recitational* items were contributed by the most talkative pupil. As against this, a median of 16½ per cent of all *self-initiated and self-assertive* contributions was made by the most talkative pupil.

In connection with *recitational* items, the median "most loquacious" child contributed as much as the cumulative contributions of the six least talkative children in the activity groups and the five least talkative children in the control groups.

As against this, in connection with self-initiated and self-assertive contributions, the median values indicate that the "most loquacious" child in the activity groups con-

tributed as much as the cumulative contributions of the 16 least talkative children. In the control classes the corresponding value was 13.

In both the activity and the control groups the median percent of contributions made by pupils in the highest quartile (in terms of amount of talking) represented approximately two-thirds of all contributions made by the class as a whole. The findings just cited were borne out to varying degrees by the individual classes.

Through the use of children's scores on various tests of information and achievement, an inquiry was made into relationships between the amount of talking children did in class and the amount of information they possessed in various areas. In general the correlations between frequency of comments of a self-initiated, self-assertive etc. variety and achievement in tests of ability to apply generalizations, of knowledge of current affairs, of working skills in the social studies, of achievement in various subject-matter skills (as measured by the Modern School Achievement tests) were positive but low.

In the activity groups there was a positive relation (represented by a median coefficient of .37) between frequency of self-assertive, self-initiated contributions and knowledge of current affairs, and a median correlation of .36 between the frequency of such contributions and cumulative scores in tests of the Modern School Achievement battery.

The other correlations, including correlations between frequency of contributions and scores on a personality test, ranged lower than this.

The correlations between achievement and frequency of contributions tended to be slightly higher in the activity than in the control groups. In both groups, however, it appears that the extent of the child's loquacity and his ability to get a hearing is primarily determined by factors other than knowledge and information such as was measured by the tests administered in this study.

Discussion

The findings in this study are consistent with the findings of an earlier study which indicate that there is more equality of pupil participation when the circumstances involve a considerable amount of direction and control by the teacher than in situations in which pupils have more freedom and self-direction.

These findings cannot be construed as a criticism of the activity program since the tendency on the part of a few children to preëmpt a large proportion of the talk appears also in the control classes. Rather, the findings emphasize a problem that arises in the management of class groups under almost any system of education. It is a problem that many teachers are aware of and have sought means to solve.

It might, of course, be argued that it is no more than proper that children who are so disposed should be free to monopolize a large share of class discussion. As against this, however, certain propositions may be raised. Many factors seem to contribute to disproportions in amount of pupil participation. There are pupils, for example, who participate little but who apparently would like to participate more. Some pupils, for example, are slower in taking advantage of an opening or shrink from interrupting others or hesitate to drown out the voice of another pupil who already is talking. On the other hand, there are youngsters who show no such compunctions. In individual cases it also appeared, from an earlier study, that extreme verbosity in the classroom situation was associated with evidences of personality difficulties. It must also be said, however, that some individual children who seldom volunteered to talk, or who spoke only briefly when they were given the floor, seemed to be content to let other youngsters carry the burden of talk.

Another point may be mentioned. One aim in connection with the "newer" procedures in education is to encourage pupils to join with the teacher in planning their work and in making decisions that affect the activities of the class. One of the difficulties in connection with this objective is to obtain genuine expressions of the interests of the group. What sometimes seems to happen is that "group decisions" which determine the choice of a given unit of work, or the procedures that will be followed in connection with a given project, actually do not reflect the sentiment of the group but are determined by a few children who happen to be very articulate, who establish priority for a certain suggestion, or by sheer volume of talk prevail upon the class as a whole to adopt

their proposals. Needless to say, the same condition often prevails among adults: The most vocal adult at a business meeting or forum, who beats all others to the draw in making a motion or in nominating a candidate for office, or in presenting a personal viewpoint, frequently brings about a decision which might not have been reached if all members of the group could have voiced their silent wishes in the matter.

Although the present investigation does not include a systematic study of means by which this problem might be solved, the writer has been able to obtain some incidental bits of information from teachers in various situations. There are teachers who are very conscious of this situation and who make deliberate efforts to promote more equitable participation by means such as a policy of rotating the chairman or leader in class undertakings and in the work of committees, by the procedure of being on the alert to give the least loquacious child a chance to be heard, and by a policy of encouraging pupils themselves to discuss the problem and to take a hand in solving it. It is likely that a policy of utilizing discipline that arises from within the group itself could be applied considerably more than it is.

At the other extreme, however, there are occasional teachers who condone the tendency of some children to preëmpt or to govern class discussion on the theory that youngsters who are disposed to take command in this way may possibly be regarded as future leaders and should have a chance to exercise their powers.

If the pupils who are encouraged to make use of privileges of this sort also happen to be able, wise and sensitive to the wishes of the group as a whole, such a policy might perhaps be defended. It also has its dangers, however. For one thing it obviously is difficult to detect potential leaders. The choice may be quite fortuitous. The child who happens to be somewhat forward or inconsiderate may win a commanding place when certain other less invasive children possess more actual or potential merit.

The findings and observations reviewed above describe a state of affairs which presents a problem in schools both of the conventional and of the newer type. The phenomena here described have implications both at the childhood and the adult level. As far as the management of the elementary school program is concerned, these phenomena indicate the need for further study of improved ways of promoting democratic procedures.

Bibliography

Below is an alphabetical list of other studies directly or indirectly related to the evaluation project of which the present investigation was a part.

1. Jersild, A. T., Thorndike, R. L., Goldman, B., and Loftus, J. J. "An Evalution of Aspects of the Activity Program in the New York City Public Elementary Schools," *Journal of Experimental Education*, VIII, (December, 1939), pp. 166–207.
2. Jersild, A. T., Thorndike, R. L., Goldman, B., Wrightstone, J. Wayne, and Loftus, John J. "A Further Comparison of Pupils in 'Activity' and 'Non-Activity' Schools," *Journal of Experimental Education*, IX, (June, 1941), pp. 303–322.
3. Jersild, A. T., Goldman, B., and Jersild, C. L. "Studies of Elementary School Classes in Action, I. A Comparative Study of the Daily Occupations of Pupils in 'Activity' and 'Non-Activity' Schools," *Journal of Experimental Education*, IX, 1941, (June, 1941), pp. 295–302.
4. Jersild, A. T., Goldman, B., and Loftus, J. J. "A Comparative Study of the Worries of Children in Two School Situations." *Journal of Experimental Education*, IX, (June, 1941), pp. 323–326.
5. Jersild, A. T., Meigs, M. F., and Brown, L. S. "A Study of the Elementary Classrooms in Action." (1939) Unpublished.
6. Jersild, A. T. "A Study by the Interview Technique of the Reactions and Appraisals of Pupils Who Have Transferred from Conventional to 'Activity' Schools." Unpublished.
7. Loftus, John J. "New York City's Large-Scale Experimentation with an Activity Program: A Metropolitan School System Takes Stock," *Progressive Education*, (1940), pp. 116–124.
8. McCall, William A., Herring, John P., and Loftus, John J. "Measuring Achievement in Activity and Control Schools in New York City," *Teachers College Record, XXXIX* (February, 1938), pp. 423–431.

9. McCall, William A., Herring, John P., and Loftus, John J. "Measuring the Amount of Activity Education in Activity and Control Schools in New York City," *Teachers College Record,* XXXIX (December, 1937), pp. 230–240.

10. McCall, William A., and Loftus, John J. "America's Largest City Experiment with a Crucial Education Problem," *Teachers College Record,* XXXVIII (April, 1937), pp. 602–606.

11. Sells, Saul B., Loftus, John J., and Herbert, Louis. "Evaluative Studies of the Activity Program of the New York City Public Schools: A Preliminary Report," *Journal of Experimental Education,* IX, (June, 1941), pp. 310–322.

12. Thorndike, R. L., Loftus, John J., and Goldman, B. "Observations of the Behavior of Children in Activity and Control Schools," *Journal of Experimental Education,* December, 1941.

13. Thorndike, R. L., Loftus, John J., and Goldman, B. "Observations of Excursions in Activity and Control Schools," *Journal of Experimental Education,* December, 1941.

14. Wrightstone, J. Wayne, Rechetnick, Joseph, McCall, William A., and Loftus, John J. "Measuring Intellectual and Dynamic Factors in Activity and Control Schools in New York City," *Teachers College Record,* XXXIX (December, 1938), pp. 237–244.

15. Wrightstone, J. Wayne, Rechetnick, Joseph, McCall, William A., and Loftus, John J. "Measuring Social Performance Factors in Activity and Control Schools in New York City," *Teachers College Record,* XL (February, 1939), pp. 423–432.

16. Wrightstone, J. Wayne, Rechetnick, J., and Loftus, J. J. "Results of Comprehensive Achievement Test in Controlled Experiment," *Teachers College Record,* XXXIX (February, 1938), pp. 431–432.

OBSERVATIONS OF THE BEHAVIOR OF CHILDREN IN ACTIVITY AND CONTROL SCHOOLS

ROBERT L. THORNDIKE
Teachers College, Columbia University
JOHN J. LOFTUS
Board of Education, New York City Public Schools
and
BERNARD GOLDMAN
Teachers College, Columbia University

This article reports certain further data on the effects of the "activity" program in the New York City public schools.* The character of the New York City experiment in curriculum revision has already been described in earlier publications[1] and will not be elaborated here. It will suffice to say that in the experimental group an effort was made to organize learning around centers of pupil activity and interest, breaking away in some measure from the prescribed subject-matter curriculum. From the schools making this change, certain ones were selected and paired with control schools for study of a variety of outcomes. Previous publications include the results of several sorts of tests and observations.

One type of data which was gathered extensively during five semesters of the experiment was classroom observation of various categories of behavior. The procedure was for an observer to spend periods of a half-hour in the classroom, keeping watch over the whole class. Whenever the observer noticed an item of behavior which seemed to him to fall under one of seven defined categories (cooperative, critical, experimental, leadership, recitational, self-initiated, and negative work-spirit behaviors), he recorded a notation in code on his record sheet. At other times, observers recorded in anecdotal form descriptions of items of behavior which they had observed. The coded observations yielded data on the frequency of occurrence of these categories of behavior in experimental and control classes, and the anecdotal records made it possible to obtain in addition some estimate of the quality of the behavior items.

In the observations described above, the observer was called upon to observe a whole class at one time. Though the observations may provide a fairly satisfactory over-all picture of the class as a whole, insofar as the categories under consideration are concerned, it seems likely that they give a very incomplete picture of the whole range of activities of any one child. It seemed worthwhile, therefore, to supplement the observations of class units with observations of individual pupils. It is with these observations of individual pupils that this paper is concerned.

These observations were carried out during the spring and winter semesters of 1940 in 4 activity and 4 control schools. Each activity school was paired with a control school from the same neighborhood, in an effort to work with populations of comparable ethnic and socio-economic backgrounds. At the beginning of each semester, 4 pairs of classes were selected in each pair of schools (making a total of 16 pairs of classes), classes being matched according to grade and according to mental ability insofar as that was known. In each class, the observer selected 5 children (making a total of 80 activity children and 80 control children) for observation. In order that the teacher might have as little opportunity as possible to bias the results, she was not informed which children in her class were being observed. The number of children in each group finally included in the study

* This study was conducted with the aid of workers affiliated with the Work Projects Administration, Research and Records Program District Office, Research Survey and Federal Data Section, Official Project No. 65-1-97-21, WP 10 Sub. 1.

[1] Arthur T. Jersild, Robert L. Thorndike, Bernard Goldman, and John J. Loftus. "An Evaluation of Aspects of the Activity Program in the New York City Public Elementary Schools." *Journal of Experimental Education*, VIII (December, 1939), pp. 166–207.

each semester was somewhat over 80 because of transfers and replacements.

The children were selected after a period or two of general observation in the classroom. No definite rules for selection were established, but it was suggested that observers select (a) children who were fairly typical of the class—neither outstandingly good nor outstandingly bad, neither outstandingly active in participation nor outstandingly passive, (b) children from different parts of the room, (c) children who were located where they could be observed readily and inconspicuously, (d) about equal numbers of boys and girls.

Data upon the 16 classes and 80-odd children in each group each semester are presented in Tables I, II, and III.

As far as grade in school is concerned, the activity and control groups are well equated. In terms of level of ability, however, it appears that our matching has been less satisfactory. It was not possible to obtain test records for all children, and the gaps in our test data were concentrated in the slow control classes, so that evidence concerning intelligence is less perfect than we might wish.

TABLE I

GRADE GROUPS STUDIED

	Spring 1940		Winter 1940	
	Activity	Control	Activity	Control
3A		1	1	1
3B	2	2	3	3
4A	3	1	1	1
4B	2	3	2	2
5A	2	2	2	2
5B	2	2	3	3
6A	3	3	4	4
6B	2	2		

TABLE II

CHARACTERIZATION OF GROUP BY SCHOOL AUTHORITIES

	Spring 1940		Winter 1940	
	Activity	Control	Activity	Control
Bright or bright average	6	6	6	7
Average	4	5	5	5
Slow average or slow	2	5	1	4
Unspecified*	4		4	

* These groups represent a complete cross-section of each grade in one school in quite a poor district. They probably correspond most nearly to slow average in intellectual level.

An examination of Table III suggests, however, that the control groups were slightly higher in mental ability, and varied rather more from class to class. Taking into account all the data on hand, it is estimated that the control classes averaged about 5 points higher in I. Q. This may handicap the activity groups slightly in certain comparisons.

PROCEDURE

In general, the procedure followed was to observe each individual pupil for a number of periods of five minutes each. A diary record was kept of the actions of the pupil during the five minutes. At first, the observations were quite exploratory, and the observers recorded whatever they saw, with little instruction or guidance. As the study progressed, after examination of the early records and discussions between the observers and Dr. Thorndike, the attention of observers was directed to certain features of the situation of which they had been less aware before and a more systematic watch was kept upon certain features of the situation which had not been so generally considered in the early reports. (For example, the responses of the teacher or of other children to the child, failures to respond on the part of the child, etc.) The median number of observation periods for a single child was 25, and these were spread out over the semester.

During the first semester, sample records were studied by Dr. Thorndike, and a set of categories was devised for classifying the items of behavior reported in the diary records. This set of categories went through several revisions. Each time it was revised, it was tried out by the observers on a sample of records, to see how well the recorded behaviors fitted into the framework of categories. Where difficulties were encountered, the set of categories was extended or revised to take care of them. Eventually a set of categories was developed which seemed to cover fairly completely the behaviors reported in the diary records. During the second semester of the experiment, these categories were used by the observers in the classrooms, and the behavior of the child was recorded directly in the code number which designated that behavior. The records for the first semester were translated into the set of behavior categories after the conclusion of the semester of observation.

TABLE III

INTELLIGENCE TEST DATA

Spring 1940

Type of Class	Activity Total N	Activity N Tested	Activity Mean I. Q.	Control Total N	Control N Tested	Control Mean I. Q.
Bright or bright average	30	30	116	29	25	130
Average	20	16	110	25	23	114
Slow average or slow	10	8	105	26	10	100
Unspecified	20	19	101			

Winter 1940

	Activity Total N	Activity N Tested	Activity Mean I. Q.	Control Total N	Control N Tested	Control Mean I. Q.
Bright or bright average	31	31	117	37	28	128
Average	27	27	114	26	20	110
Slow average or slow	5	0	—	20	3	100
Unspecified	20	16	105			

In making a statistical analysis of the records for a given child, each five minute period was taken as a unit of observation and the rater determined which categories of behavior had occurred for that child during that period. No account was taken of repetitions of the same behavior during a single period of duration of the behavior; it did not seem practical to do so. The record for a given period of observation consisted, at this point, of a list of the categories of behavior which had, in the first semester, been found in the diary record or, in the second semester, been recorded by the observer for that period.

The reporting of behavior items or classification of the items in a diary record is necessarily somewhat subjective. It is apropos to ask at this time how well observers agreed in the categories of behavior which they saw in the same records. Data on this point were gotten by having all four observers categorize the behaviors from a sample of the records. It was then possible to determine the percent of agreement between pairs of individuals analysing the same records. If a category was reported as present in the diary of a particular five-minute behavior sample by both observers of a pair, an agreement was recorded. If it was reported as present by one but not reported by the other, a disagreement was recorded. The percent of agreement between two observers was determined by the formula

$$\% = \frac{\text{Agreements}}{\text{Agreements} + \text{Disagreements}}$$

Using this formula, the percent of agreement was determined for each pair of observers. These percents are over-all values including all the different categories of behavior. The percents are as follows:

Observers	Percent Agreement
A vs. B	46
A vs. C	53
A vs. D	48
B vs. C	52
B vs. D	50
C vs. D	53

We see that, in general, agreement clusters around 50 percent. This is disappointingly low. An appearance of greater reliability could be gotten by using the formula

$$\frac{2 \times \text{Agreements}}{2 \times \text{Agreements} + \text{Disagreements}}$$

that has rather conventionally been used in reporting the reliability of observation studies.* The percent of agreement by this method, which is more directly comparable to the percent of agreement from other observation studies, would be about 67. This still does not impress one as particularly high.

In part the low agreement is due to a fine subdivision of categories which makes minor disagreements possible. It was found, for example, that the average agreement for several sub-categories of recitation was 61%, while if these were combined into a single "recitation" category, the figure was raised to 73%. Similarly, several sub-categories of self-initiated activity gave an average agree-

* Arrington, R. E. *Interrelations in the Behavior of Young Children*. Child Development Monographs, No. 8, Teachers College, Columbia University, Bureau of Publications, 1932.

ment of 16%, while a combined category yielded an agreement value of 21%. However, it is clear that even the combined categories do not yield a high percent of agreement, and this lack of reliability reduces the faith which we can have in the results from these observations. The results must be viewed as suggestive and supplementary rather than conclusive.

Analysis of Results

We turn now to the comparison of the activity and control groups in the various categories of behavior. In analysing the results, the first step was to obtain a score for each individual for each of the behavior categories. This score was the percent of observation periods in which that behavior occurred (according to the classification of behavior items by the rater). For example, if a child's record showed him to respond to direct teacher questioning in 5 out of 25 observation periods, his score would be 20. After a score has been computed for each individual for a particular category, a frequency distribution of the scores can be made for the activity group and for the control group. From these frequency distributions it is possible to compute the mean and standard deviation for each group. The standard error of the mean, standard error of the difference between activity and control means, and critical ratio are then determined. Reference to these critical ratios aids us in judging which differences are too large to be explained as chance fluctuations in samples drawn from the same universe. No standard errors or critical ratios are reported for categories in which the average score is very small (less than 1.00).

In Table IV are reported the activity and control group means and the difference between them for each semester. Differences significant at the 5 percent level are marked with a single asterisk, and those significant at the one percent level with a double asterisk.

An examination of Table IV shows a number of differences which are so large that they could hardly have arisen by chance alone. In spite of this, however, there seems more of similarity than of difference in the pattern of behaviors in the two school situations. If, for example, we list the 10 most frequent behaviors for the activity group and the control group for each semester, we find seven categories which appear in all four lists and two more which appear in three out of the four. Listening to teacher or other pupil, carrying out routine instructions, and fidgeting are, in that order, the three behaviors most frequently reported in both groups. There are seven categories of behavior for which a difference as great as five in percent of periods showing the behavior is obtained for the average of the two semesters. These are:

(1) Under general stimulus from the teacher, carrying out an individual activity of making or producing.
(2) Initiating communication with others.
(3) In response to a question asked the class by the teacher, volunteering and being permitted to respond.
(4) As above, only not being permitted to respond.
(5) Working as an individual studying an assigned exercise.
(6) Observing or passively attending to teacher explanation, student recitation or report, question-answer, etc.
(7) Carrying out mechanical detail of class routine.

The first two occurred more frequently in the activity classes, and the last five more frequently in the control classes. Of 56 categories, 30 showed differences in average frequency of occurrence of less than one percent, when results from the two semesters were combined.

Now let us examine the results more in detail. We will consider each major category in turn. Our remarks here will generally be rather tentative, because most of the differences are too small to be conclusive.

The first category has to do with class recitation and discussion. It seems clear that there was much more teacher questioning and response thereto in the control group. The results for the control group show almost twice as high a score in the various subcategories of teacher questioning. This is in agreement with earlier observation of the class group as a unit, which consistently showed more "recitational activity" in the control classes. At the same time, however, the amount of recitation in response to teacher question remains substantial in the activity group.

TABLE IV

AVERAGE PERCENT OF PERIODS IN WHICH CERTAIN BEHAVIOR CATEGORIES APPEARED

Category	Spring 1940 Average Activity	Spring 1940 Average Control	Spring 1940 Diff.	Winter 1940 Average Activity	Winter 1940 Average Control	Winter 1940 Diff.
I. Contribution to recitation or discussion involving part or all of class						
A. When question is asked **him** directly by teacher						
(1) responds	3.86	6.38	—2.52**	5.68	8.96	—3.28**
(2) fails to respond	.53	.97	—0.44	.17	.26	—0.09
B. When question is asked **class or group** by teacher						
(1) volunteers and is permitted to respond	4.02	10.32	—6.30**	9.36	13.12	—3.96*
(2) Volunteers, but not permitted to respond	14.80	24.40	—9.62**	12.24	23.15	—10.91**
(3) does not volunteer to respond	6.93	8.67	—1.74	8.16	14.25	—6.09**
(4) responds with class as a group	3.46	5.22	—1.76*	2.66	4.64	—1.98**
(5) insistent in effort to respond	3.90	5.72	—1.82*	3.02	5.84	—2.82**
C. When question is asked class or group by **pupil chairman**						
(1) volunteers and is permitted to respond	0.68	0.34	0.34	1.80	0.32	1.48
(2) volunteers, but not permitted to respond	1.82	1.28	0.54	1.25	2.08	—0.83*
(3) does not volunteer to respond	2.67	0.86	1.81**	1.51	1.28	0.23
(4) Responds with class as a group	0.32	0.33	—0.01	0.09	0.00	0.09
(5) insistent in effort to respond	0.79	0.88	—0.09	0.13	0.45	—0.32
D. Spontaneously volunteers contribution to class						
(1) is permitted to make contribution	4.62	3.48	1.14	4.60	3.81	0.79
(2) not permitted to make contribution	4.22	4.80	—0.58	3.14	3.20	—0.06
E. Presents prepared report to class	1.41	1.09	0.32	3.40	1.90	1.50**
II. Working as an individual carrying out definitely assigned exercise						
(1) Study	15.40	17.85	—2.45	8.02	26.71	—18.69**
(2) Test or exercise	21.04	22.50	—1.46	21.75	22.28	—0.53
(3) Production-writing, drawing	10.74	7.76	2.98**	9.82	6.74	3.08**
(4) Singing or choral reading	2.98	4.44	—1.46	4.48	2.92	1.56*
(5) Physical activity	1.73	3.48	—1.75**	1.61	2.70	—1.09**
III. Teacher-originated and approved activity, but without specific and immediate direction						
A. Participates actively with others	5.28	3.63	1.65	7.28	2.36	4.92**
B. Carrying out activity by self						
(1) Study	2.19	2.88	—0.69	8.55	3.88	4.67**
(2) Test or exercise	0.77	0.51	0.26	6.36	3.42	2.94**
(3) Production	6.92	1.82	5.10**	10.94	4.20	6.74**

TABLE IV—Continued

AVERAGE PERCENT OF PERIODS IN WHICH CERTAIN BEHAVIOR CATEGORIES APPEARED

Category	Spring 1940			Winter 1940		
	Average Activity	Average Control	Diff.	Average Activity	Average Control	Diff.
IV. Self-originated and directed activity						
A. Clearly related to present activity of class	4.08	2.61	1.47	1.17	1.01	0.16
B. Not related to class activity but acceptable	4.05	3.86	0.19	6.41	3.90	2.51**
C. Clearly unacceptable to teacher and class purpose	5.73	5.62	0.11	4.12	3.26	0.86
D. Relation to work of class indeterminate	3.20	1.99	1.21*	3.75	1.98	1.77**
E. Activity involves participation with others	1.19	0.90	0.29	0.25	0.00	0.25
F. Recipient of "outlaw" behavior of another	0.47	0.37	0.10	0.13	0.17	—0.04
V. Observes or passively attends to some activity						
A. Audio-visual material	2.18	1.93	0.25	3.30	1.98	1.32*
B. Teacher explanation, student recitation or report, question-answer, etc.	59.60	64.44	—4.84	85.10	92.24	—7.14
C. Activity of small group or of individual						
(1) Legitimate activities	5.04	5.48	—0.44	5.49	3.58	1.91*
(2) "Outlaw" activities	2.42	1.17	1.25*	1.88	1.40	0.48
(3) Activities unspecified	6.09	4.30	1.79	4.56	5.82	—1.26
VI. Cooperates in carrying out mechanical detail of class routine	37.87	44.56	—6.69*	34.05	41.06	—7.01**
VII. Gives or receives cooperation						
A. Offers or gives help	12.75	9.99	2.76	11.58	7.14	4.44**
B. Refuses to give help, when requested	0.75	0.24	0.51	0.05	0.05	0.00
C. Solicits help	8.18	4.84	3.34**	6.60	5.20	1.58
(1) Takes help without consent						
(a) Takes material or supplies	0.99	0.86	0.13	0.33	0.13	0.20
(b) Observes and copies work	1.22	0.49	0.73	1.94	0.82	1.12
D. Help is offered to child observed						
(1) Accepted by him	1.30	1.02	0.28	2.16	1.76	0.40
(2) Rejected by him	0.38	0.25	0.13	0.10	0.09	0.01
VIII. Communicates with another pupil						
A. Initiates communication	34.09	15.72	18.37**	28.30	18.96	9.34**
B. Receives or responds to communication	8.20	6.52	1.68	11.86	6.92	4.94**
IX. Appears inactive and inattentive—day dreams	7.38	4.56	2.82**	4.72	7.32	—2.60*
X. Fidgets	35.33	36.81	—1.48	23.05	29.96	—6.91*
XI. Displays feeling						
A. Expresses enjoyment, pleasure, or enthusiasm	17.25	16.98	0.27	16.32	18.96	—2.64
B. Expresses annoyance, displeasure	3.69	1.85	1.84**	1.64	1.04	0.60

TABLE IV—Continued

AVERAGE PERCENT OF PERIODS IN WHICH CERTAIN BEHAVIOR CATEGORIES APPEARED

Category	Spring 1940 Average Activity	Spring 1940 Average Control	Spring 1940 Diff.	Winter 1940 Average Activity	Winter 1940 Average Control	Winter 1940 Diff.
XII. Disciplines or is disciplined						
A. Reprimands other child, or criticizes conduct	1.76	0.56	1.20	0.40	0.21	0.19
B. Is reprimanded						
(1) by teacher	3.82	4.14	—0.32	4.06	4.62	—0.56
(2) by other child	0.87	0.38	0.49	1.20	1.16	0.04
XIII. Commends or is commended						
A. Approves action of other child or teacher	0.78	0.43	0.35	0.62	0.09	0.53
B. Is approved						
(1) by teacher	1.20	2.45	—1.25**	2.60	2.78	—0.18
(2) by another child	0.18	0.74	—0.56	0.67	0.39	0.28

There was a somewhat higher frequency of discussion and questioning in which a pupil is the chairman in the case of the activity classes. However, this situation was reported quite infrequently even in these groups. There is some suggestion that the activity children responded less actively and enthusiastically to this situation when it occurred. In frequency of spontaneous contributions we find no significant differences. There is some evidence for more use of prepared pupil reports in the activity classes.

Individual work upon a teacher-assigned task is frequently reported in both activity and control classes. Tests and exercises seem to be about equally common, but productive activities, such as preparing drawings, compositions, craft work, etc., are observed more often in the activity groups. This is in accord with data upon semi-independent production presented further along in the table, and both of these find confirmation in the more frequent "experimental activity" reported in earlier studies of the class group as a whole. This difference in frequency of both immediately and remotely teacher-directed productive work appears as one of the fairly clear differences between our two groups. Especially in the winter semester, the control groups show more study on definite teacher-assigned tasks. This contrasts with partially self-directed study, which is reported more frequently in the activity classes. Twice as much group calisthenics and organized physical activity is reported in the control group, suggesting, perhaps, the greater formality of that situation and the greater need for an authorized opportunity to relax.

In the categories involving semi-self-initiated and self-initiated activities, greater frequencies are quite uniformly reported for the activity group but many of the differences are small. Reliable differences are found for both semesters in semi-self-directed production, and for the winter semester in semi-self-directed group activity, individual study, individual test or exercise, and independent activity of an apparently acceptable or at least neutral character. The direction of the differences in these categories are in conformity with classroom observations on the frequency of self-initiated activities in the two types of classes.

It is clear that in both groups a great deal of time—with perhaps a little more in the control groups—is spent in following, with an unspecified degree of attention and interest, the activities of the teacher and/or other pupils. This is the most common category for both activity and control groups. Most of this watching and listening is directed toward legitimate school activities, but there is a suggestion that "outlaw" activities may be more frequently observed in activity groups. Another oft-recurring event, appearing again and again with both groups, is carrying out mechanical details of classroom routine—cleaning desks, passing paper, reporting scores, etc., etc. The mechanics of running the class bulks quite large in the picture in both groups, though the frequencies in the control group are higher.

Cooperation appeared more frequently among the activity children, though most of the differences are not large. Interesting is the fact that these children appeared not

only more ready to give help but also more ready to solicit it (and even take it without asking). These differences are also in line with earlier evidence on cooperative activities obtained from observation of whole-class units.

One of the most conspicuous differences between the two groups is in frequency of communication. The activity children appear to be much freer in communicating with their fellows. Apparently the situation in which they find themselves makes informal give and take among pupils a good deal more possible. This finding, together with much of the rest of the material which we are able to present here, provides more of a description than of a direct evaluation of the activity program.

The data on inactivity and day-dreaming are inconsistent for the two semesters. Fidgeting was reported often in both groups—perhaps a little more so in the control. The standard for reporting fidgeting was quite low, so that the frequent occurrence of fidgeting should probably not be a matter of serious concern to us.

With regard to display of feeling, it is interesting that pleasant, positive reactions greatly outnumber the negative ones in both cases. However, we may note that expressions of annoyance and displeasure are somewhat more prevalent in the activity group than in the control. Whether frustrations are more frequent in the former situation, or only more freely expressed, is not indicated in the data. During the spring semester, it appeared that the ratio of reprimands to approvals was higher in the activity than in the control classes, but this is not supported by data from the winter semester.

In summary, there is a good deal of similarity in the frequency with which behavior categories are reported in the activity and control classes. The control classes show a clear picture of more teacher question and pupil answer. They probably also involve more (1) assigned study (2) listening and watching the teacher and other pupils (3) routine of class management. The activity classes show (1) more student chairmanship, (2) more productive work in writing, art, etc., both directly and indirectly teacher sponsored, (3) a general picture of somewhat more semi-independent and independent activity, (4) more giving and taking of help, (5) more communication between pupils. Other differences are small or contradictory.

OBSERVATIONS OF EXCURSIONS IN ACTIVITY AND CONTROL SCHOOLS

ROBERT L. THORNDIKE
Teachers College, Columbia University
JOHN J. LOFTUS
Board of Education, New York City Public Schools
and
BERNARD GOLDMAN
Teachers College, Columbia University

In the evaluation of the activity program in the New York City elementary schools,[1] the data so far reported have been based upon either classroom observations or paper-and-pencil tests and questionnaires. It seemed likely that significant evidence concerning the outcomes of the program might be obtained if pupils could be observed in situations other than those arising within the limits of the classroom. The question has been raised as to whether the differences in classroom behavior reported in the earlier evaluation of the program may not reflect the temporary classroom situation rather than any more general and permanent changes in the pupils. We might ask, for example, whether the more frequent evidences of initiative shown by the activity pupils are limited to the classroom situation, or whether they represent gains which will also appear in out-of-class activities. Observation of trips to points outside the school building seemed to provide one means for getting information upon this transfer.[2]

At the time of this study (1940–41), a WPA project supplied personnel and made the arrangements for a large number of trips from schools to places of interest in and around New York City.[3] In the typical case, the trip was conducted by a WPA guide, with the teacher also present and playing some part in the direction and conduct of the expedition. Parks, museums, business houses, historic sites, and other points of interest served as the goals of these excursions. The greater number of excursions took a full school day.

The material to be reported here is based upon diary accounts of these trips and upon ratings of various aspects of behavior during the trip. The diary record was made by an observer, who accompanied the class on the expedition and made running notes of what happened. After a number of diary records had been made, these were examined by the directors of the study to see what form the analysis of the data might take. On the basis of this examination, several scales were prepared, to be used in rating different features of behavior. From that time on, the observers kept diary records as before, but also rated the group upon the several scales. In the case of trips made before the rating scales had been prepared, ratings were subsequently made, based upon the diary records.

In practice, it appeared that the trips were something less than ideal in providing a free non-school situation for observation. The children were accompanied by an adult guide and by the teacher, and were rarely out of sight of one or both of these adults. Concern for the safety of the children and for the comfort of the public with whom they came in contact gave rise to rather close supervision by those in charge of the group. It is the impression of the authors that for some groups at least more freedom and spontaneity was permitted in the classrooms than on the trips. Be that as it may, the trips

[1] For a discussion of the activity program and data on other aspects of the program of evaluation the reader is directed to: A. T. Jersild, R. L. Thorndike, B. Goldman, and J. J. Loftus, "An Evaluation of Aspects of the Activity Program in the New York City Public Elementary Schools," *J. Exper. Educ.*, VIII, (1939) 166–207.

[2] This study was conducted with the aid of workers of the Work Projects Administration, Research and Records Program District Office, Research Survey and Federal Data Section, Official Project No. 65-1-97-21 WP 10 Sub. 1.

[3] This service is under the direction of Mr. Orestes La Polla, whose cooperation in this study is gratefully acknowledged.

did provide an opportunity to observe children in a *different* situation from that of the classroom, even though that situation was not entirely free from the direction and control of the school.

In evaluating the trip, the observer tried to rate (either directly at the time of the trip, or indirectly through the diary record) twelve features of trip behavior. It was not always possible to give a rating upon some of the scales, because the trip had provided no opportunity to observe the behavior in question. Several scales, for example, referred to behavior during the lunch period. If the trip did not include a lunch period (or if no mention of behavior during lunch was included in the diary record), obviously no rating of luncheon behavior was possible. The characteristics of the trip to be rated, on a scale from 1 to 9, were as follows:

A. *En route to and from the museum, park, etc.*

1. *Rigidity of adult control.* Amount of regimentation and direction by the teacher or guide, admonitions to stay in line, keep quiet, etc.

2. *Social acceptability of behavior.* Amount of behavior which would be deemed socially objectionable in that situation by a tolerant adult.

3. *Quality of contacts with the public.* (In many cases, the observers reported "No evidence upon which to judge" for this scale, and so records are quite incomplete.) Extent to which behavior towards the public tends to leave generally negative and unpleasant impression.

B. *Seeing exhibits, hearing lecture, etc.* (the instructional part of the trip.)

1. *Rigidity of imposed control.* As above, except that the types of restraints may differ somewhat.

2. *Social acceptability of behavior.* As above, except that the standard of judgment must be varied for different situations.

3. *Quality of interest and participation.* Activeness of interest and its spread through the group.

C. *Lunch period.*

1. *Social spontaneity.* Amount of active social interaction during the luncheon period.

2. *Acceptability of luncheon behavior.* Amount of behavior of disruptive and disturbing character.

3. *Responsibility for clean-up.* Amount of cooperation in cleaning up debris.

D. *Art or construction period.* (These occurred for only a few groups.)

1. *Quality of interest and participation.* As in B3, taking account of the differences in the character of the situation.

E. *Play period.* (This, again, occurred for only a few groups.)

1. *Completeness of participation.* Unanimity of active participation.

2. *Level of good-will.* Amount of dissension and conflict.

In all, 41 trips from activity schools and 77 trips from control schools were observed. As the assignments for observation were made on short notice and depended upon the trips which were available, it was not possible to make any systematic matching of control and activity groups. However, there is no reason to believe that there was any systematic bias introduced in favor of either the control or activity schools. Whatever differences appear in the constitution of the groups may well represent chance fluctuations in a fairly small sample of classes.

Some information about the groups is available and this is assembled in Table I. The information is incomplete and sketchy, but it presents a picture of reasonable balance in grade, socio-economic status, and ability in the two populations. The most

TABLE I

CHARACTERISTICS OF ACTIVITY AND CONTROL GROUPS OBSERVED ON OUT-OF-SCHOOL TRIPS

	Number of Classes	
	Activity	Control
A. Grade		
Ungraded	1	4
First	1	1
Second	0	3
Third	2	8
Fourth	10	13
Fifth	12	23
Sixth	15	25
B. Racial group		
Negro	1	13
Negro and White	2	5
White	15	26
Unspecified (probably white)	23	33
C. Socio-economic status		
Average	7	14
Low	10	20
Unspecified	24	43
D. Intelligence		
High	4	13
Average	8	14
Low	5	9
Unspecified	24	41
E. Median Class Size	28.4	29.2
F. Destination of trip		
Museum	18	33
Airport	4	4
Park	3	12
Downtown district	12	19
Other	4	9

striking difference is in the proportion of Negro groups reported, there being 1 group in the 41 activity classes and 13 among the 77 control classes. However, even when these Negro groups are included, the evidence which we have does not suggest that the control groups have a large proportion of low intelligence or low socio-economic level. Results are reported both for the total population and for the population excluding the Negro classes. The inclusion or exclusion of these groups makes only a small difference in the results.

The essential data are the average rating for activity classes and for control classes on each of the scales which were described earlier. Tables II and III include, for each scale, the mean rating for activity and control classes, the difference between the two means, the standard error of the difference, and the critical ratio. The number of classes for which ratings were available is reported for activity and control groups also. Table II includes the Negro classes, while Table III does not. In these tables some scales have been reversed, in order that each scale might be expressed in terms of the "desirable" quality. The scales have been so arranged that 1 always represents the most desirable and 9 the least desirable behavior. "Desirable" is interpreted as freedom from control, social acceptability, active interest, responsibility for clean-up, and a high level of social participation. The difference column reports the control mean minus the activity mean, and when this is positive, the difference is to be interpreted as favoring the control group.

As we examine the results of Tables II and III, we see certain tendencies emerging so consistently that one would doubt whether they represent simply the vagaries of chance. Though rated as having somewhat more freedom from teachers' control en route to the museum, park, or point of interest, the activity classes are reported as behaving somewhat more acceptably. The data with regard to public contacts were inconclusive, partly because no report was made on this in the case of many groups. However, a determination of the per cent of all the classes reported as having negative contacts with

TABLE II

COMPARISON OF ACTIVITY AND CONTROL CLASSES IN RATING UPON TRIP BEHAVIORS

(Total Groups)

	Activity		Control			S. D.	Critical
	N	Mean	N	Mean	Diff.*	Diff.	Ratio
A. —En route to and from the museum, park, etc.							
1. Freedom from adult control	39	4.04	75	5.25	1.21	0.39	3.1
2. Social acceptability of behavior	39	2.37	75	3.56	1.19	0.37	3.2
3. Quality of contacts with the public	11	5.55	37	5.66	0.11	0.81	0.1
B. —Seeing exhibits, hearing lectures, etc.							
1. Freedom from adult control	41	3.83	75	5.08	1.25	0.41	3.0
2. Social acceptability of behavior	41	2.20	75	3.05	0.85	0.30	2.8
3. Quality of interest and participation	41	2.11	75	2.81	0.70	0.30	2.3
C. —Lunch period							
1. Social spontaneity	35	2.90	60	3.68	0.78	0.45	1.7
2. Acceptability of luncheon behavior	33	2.56	60	3.21	0.65	0.49	1.3
3. Responsibility for clean-up	18	3.22	41	4.80	1.58	0.61	2.6
D. —Art or construction period							
1. Quality of interest and participation	8	2.25	23	3.35	1.10	0.66	1.7
E. —Play period							
1. Completeness of participation	13	2.92	18	2.17	—0.75	0.78	1.0
2. Level of good-will	12	3.33	18	3.25	—0.08	0.70	0.1

* In each case, positive difference means better score, i. e., more desirable behavior, on the part of the activity group.

TABLE III

COMPARISON OF ACTIVITY AND CONTROL CLASSES IN RATING UPON TRIP BEHAVIORS

(Negro Groups Not Included)

	Activity		Control			S. D.	Critical
	N	Mean	N	Mean	Diff.*	Diff.	Ratio
A. —En route to and from the museum, park, etc.							
1. Freedom from adult control	38	4.05	63	5.19	1.14	0.41	2.8
2. Social acceptability of behavior	38	2.38	63	3.57	1.19	0.38	3.1
3. Quality of contacts with the public	11	5.55	31	5.18	—0.37	0.81	0.5
B. —Seeing exhibits, hearing lectures, etc.							
1. Freedom from adult control	40	3.85	62	5.07	1.22	0.42	2.9
2. Social acceptability of behavior	40	2.18	62	3.14	0.92	0.31	3.0
3. Quality of interest and participation	40	2.10	62	2.72	0.62	0.31	2.0
C. —Lunch period							
1. Social spontaneity	34	2.94	48	3.50	0.56	0.46	1.2
2. Acceptability of luncheon behavior	32	2.59	47	3.31	0.62	0.51	1.2
3. Responsibility for clean-up	18	3.22	36	4.54	1.32	0.62	2.1
D. —Art or construction period							
1. Quality of interest and participation	8	2.25	22	3.13	0.88	0.65	1.4
E. —Play period							
1. Completeness of participation	13	2.92	14	2.21	—0.71	0.81	0.9
2. Level of good-will	12	3.33	14	3.36	0.03	0.79	0.0

* In each case, positive difference means better score, i. e., more desirable behavior, on the part of the activity group.

the public (rated on the negative side of the neutral point of the scale) gives 12.5% for the activity and 30% for the control groups.

At the trip destination (exhibit, lecture, etc.), the picture of less rigid control and more acceptable behavior in the activity groups was maintained. There is also some suggestion of a higher level of interest and participation in the activity groups.

The differences in rating of luncheon behavior were inconclusive, but such as they were they favored the activity classes. Too few classes were observed in either art or play situations to yield evidence of any significance.

Any conclusions must be tempered by a realization of the limitations in the data upon which they are based. In this instance, one limitation is the small and possibly poorly-matched sample of classes, and another is the probability of low reliability and the possibility of bias in the ratings of the trips. Subject to these limitations, the observations suggest that the activity program has produced certain gains in ability to maintain self-discipline in these out-of-class situations, and that possibly other gains in degree of interest and participation and in social responsibility also accrue. Certainly, these data give no support to a claim that the activity program leads to disorder and lack of self-control. The situations observed were still somewhat schoolish, and perhaps represented more familiar occurrences for the activity groups, but insofar as they carry observations beyond the classroom, they provide some support for the belief that the differences in classroom behavior observed in earlier phases of the evaluation study are maintained in out-of-class situations.

LANGUAGE DEVELOPMENT OF CHILDREN IN TWO TYPES OF RURAL ENVIRONMENTS

G. M. Worbois
Child Welfare Research Station
University of Iowa

Introduction

Although rural communities differ in many ways, there are two major patterns that may be distinguished. One type consists of a small, geometrically-determined unit in which people who happen to live within the confines of the surveyor's lines are brought together. They support a small, one-room school which the children are expected to attend. The other community pattern brings together many families living in a geographically- and socially-determined area and having closely related interests. We may call the first the one-room-school rural environment; the other the consolidated-school rural environment.

In this study children in these two different kinds of rural-school environments are studied, with the purpose of finding out how the two types of communities affect the development of the children. One group attends a rural consolidated school with about twenty to twenty-five pupils per grade, and the other attends immediately adjacent one-room schools. Several factors should be constant for these two groups (such as cultural background, racial descent, etc.,) which would differ for groups more widely separated.

There seems to be some evidence that the rural environment of the one-room schools provides inadequately for language development of the rural child as shown by Baldwin, Fillmore, and Hadley (1). The question of the present study is concerned with this problem and may be stated: "How do children in 'one-room-school environments' compare in certain aspects of language development with children of the same age and the same general ability (as measured by the IQ) in the 'consolidated-school environment?' Three aspects which have been studied are reported below.

* This study was directed at the Iowa Child Welfare Research Station by Dr. Ralph H. Ojemann.

The Tests

Socio-Economic Concept Recognition Test

Five words commonly used to designate socio-economic processes were selected (Capitalism, Conscription, Inflation, National Defense, and New Deal). Any five words might have been selected, but attempt was made to find words which were widely used to identify social, economic, and political problems of current significance. It was purposed to find out how well the subjects could apply these words to situations they commonly represent.

The "situations" were provided by pooling the written opinions of eighteen university students. These descriptions were intended to be the "common denominator" of various meanings so that intelligent and informed people would not greatly disagree in saying that the word was a proper "label" for the situation. Five competent, well-read judges with varying backgrounds agreed perfectly that each of the resulting five situations "belonged with" the word it was intended to describe, under conditions of the test. The task for the subjects then was to match the words with the situations. The subjects were the pupils in grades four and six of one consolidated school and thirteen adjacent one-room schools. The score was simply the number of words matched with the situation to which it applied.

Word Fluency Test

A test of the use of words in a "creative" way was desired. The test devised required the fluent production of language symbols in response to incomplete situations. The subjects were allowed thirty seconds to write as many words as possible to complete each of forty-five such phrases as "*through the door came a . . .*," "*ice on the . . .*," "*the distant beat of a . . .*" Preliminary testing showed that pupils for whom no time limits were given secured scores not significantly different from those secured by pupils who were

restricted to the time limit. The test was given to all pupils in grades three, four, and six of both rural groups. The reliability of the test as found by correlating odd-even items for grades three to six was r = .95 ± .005. The score was simply the number of words produced in the 22.5 minutes of the test.

Basic Language Skills Test

This test is Part C of the *Iowa Every-Pupil Tests of Basic Skills*. Again, the test was given to all pupils in grades three, four, and six of the one-room and consolidated school samples. The raw scores were changed to standard scores according to Tables 1 and 2 of *Supplement to Manual for Interpretation: Norms for Iowa Every-Pupil Tests of Basic Skills*.

The Results

Socio-Economic Concept Recognition Test

Subjects in the two groups were matched on the basis of chronological age and individual Stanford-Binet IQ. Twenty-five pairs resulted from this matching. Table 1 shows that in both age and IQ the differences are small and are not significant according to the *t*-test of Fisher. On the Socio-Economic Concept Recognition Test, however, the children in the consolidated school secured an average score of .64 higher than the average score of the one-room-school children. This difference is statistically significant at the two per cent level of confidence.

These results show that the children of the one-room schools did not understand the meaning of the words (as shown by applying them to situations they commonly represent) as well as the children of the consolidated school. That is, the children in the one-room schools agree in this respect less well with people of such intelligence and knowledge as that of the judges. It is not safe to conclude that the one-room-school children know less about the actual socio-economic processes the words designate, although research directed toward this end may be suggested by these results.

Word Fluency Test

The pupils in the consolidated- and the one-room-school samples were equated on the basis of (1) chronological age and (2) Stanford–Binet IQ by the analysis of co-variance. All the subjects in both groups taking the test were thus included in the comparison. The means on the Word Fluency Test for the two groups can be seen in Table 2. Those in the consolidated school gave on the average 25.94 more words than those in the one-room schools. This difference after allowance has been made for differences in IQ and age is statistically significant beyond the five per cent level of confidence according to Fisher's F-test.

Basic Language Skills Test

Practically the same pupils were included in this sample as were included in the Word Fluency Test, the differences being due to absences on the days the tests were given. On this test the children in the consolidated-school environment secured a mean score of 55.29 while those in the one-room schools had a mean score of 41.97 (Table 3). This represents a difference in grade placement of about a year and three months. When differences between the two groups in IQ and age have been allowed for by the analysis of co-variance, this difference in means is significant beyond the one per cent level of confidence according to the F-test.

Conclusions

The consistency with which the children in the one-room-school environment of this study appear inferior to children of equal IQ and age in the consolidated-school environ-

Table 1

Socio-Economic Concept Recognition Scores for Matched Sample of Consolidated- and One-Room-School Children

	IQ			C. A.			Score		
	N	M	σ	N	M	σ	N	M	σ
Consolidated	25	102.96	14.42	25	10.25	.98	25	2.72	1.48
One-Room	25	103.12	15.21	25	10.03	1.11	25	2.08	1.32
t-Test of Difference		.17			.22			2.40	
Level of Confidence		90 per cent			90 per cent			2 per cent	

TABLE 2

MEAN SCORES ON WORD FLUENCY TEST COMPARING CONSOLIDATED- AND ONE-ROOM-SCHOOL CHILDREN BY ANALYSIS OF CO-VARIANCE, HOLDING AGE AND IQ CONSTANT

	Consolidated	One-room	Difference	F-test	L. C.
N	64	55			
M	153.25	127.31	25.94	4.17	5 per cent
σ	46.0	59.5			

TABLE 3

MEAN SCORES ON BASIC LANGUAGE SKILLS TEST COMPARING CONSOLIDATED- AND ONE-ROOM-SCHOOL CHILDREN BY ANALYSIS OF CO-VARIANCE, HOLDING AGE AND IQ CONSTANT

	Consolidated	One-room	Difference	F-test	L. C.
N	62	61			
M	55.29	41.97	13.32	12.1	1 per cent
σ	16.5	13.8			

ment is convincing evidence that these results were not anomalies. The environment of these one-room schools appears to be associated with certain deficiencies in language development.

The three tests reported in this paper are short and preliminary to more intensive study, but even in such a short-time sampling the indication seems to be clear that this is a differential aspect in the development of rural one-room- and consolidated-school children. Whether similar results will be found in other rural areas and in other aspects of language development remains an important problem.

REFERENCES

1. Baldwin, B. T., Fillmore, Eva A. and Hadley, Lora. *Farm Children: An Investigation of Rural Child Life in Selected Areas of Iowa.* (New York: D. Appleton and Co., 1930). Pp. 337.
2. Baldwin, B. T. and Fillmore, Eva A. "The Mind of the Rural Child," *Psychological Bulletin* (1928), pp. 25, 185–186.
3. Haggerty, M. E. *Rural School Survey of New York State: Educational Achievement.* (Philadelphia: Wm. F. Fell Co., 1922). Pp. 223.
4. Henson, E. M. *Growth of Children in Consolidated and Non-Consolidated Schools,* Unpublished Master's Thesis, University of Iowa, Iowa City, 1940.
5. Kyte, G. C. "Pupil Status in the Rural Elementary School," *Thirtieth Yearbook,* National Society for the Study of Education, 1931. Pp. 272 (pp. 25–54).
6. National Education Association, Dept. of Rural Education, Report of the Committee on Program and Policy. *A Policy for Rural Education in the United States,* 1940, Pp. 47.
7. Trabue, M. R. "Are Rural Pupils Different?" *Nation's Schools* (1938), pp. 21, 42.